PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/56, 15/82, 9/44, G01N 33/53, C12P 21/08, C12Q 1/68, C07K 16/40, C08B 30/00, C08L 3/00

A1

(11) International Publication Number: **WO 99/06575**

(43) International Publication Date: 11 February 1999 (11.02.99)

(21) International Application Number: PCT/GB98/02280

(22) International Filing Date: 30 July 1998 (30.07.98)

(30) Priority Data:
9716185.5 31 July 1997 (31.07.97) GB

(71) Applicant *(for all designated States except US)*: PLANT BIOSCIENCE LIMITED [GB/GB]; Norwich Research Park, Colney Lane, Norwich NR4 7UH (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: EDWARDS, Elizabeth, Anne [GB/GB]; 4 Queens Road, Hethersett, Norfolk NR9 3DB (GB). SMITH, Alison, Mary [GB/GB]; 44 Cambridge Street, Norwich, Norfolk NR2 2BB (GB). MARTIN, Catherine, Rosemary [GB/GB]; 21 The Street, Brooke, Norwich, Norfolk NR15 1JW (GB). BUSTOS GUILLEN, Regla [ES/GB]; 347 Dereham Road, Norwich, Norfolk NR2 3UT (GB).

(74) Agents: KREMER, Simon, M. et al.; Mewburn Ellis, York House, 23 Kingsway, London WC2B 6HP (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: STARCH DEBRANCHING ENZYMES

(57) Abstract

Disclosed are isolated nucleic acids comprising nucleotide sequences which encode a polypeptide which have the properties of isoamylases, which are starch debranching enzymes, and are obtainable from Solanum tuberosum (e.g. amino acid sequence shown in any one of SEQ ID Nos. 4, 5 or 6, encoded by nucleic acid SEQ ID Nos. 1, 2, or 3). Also disclosed are variants of the same, and methods for isolating or producing these, plus also corresponding polypeptides and antibodies. Further aspects of the invention include vectors, transformed cells, and transgenic plants containing the nucleic acids of the present invention, plus also starches having modified branching characteristics and methods and materials for producing and using the same.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

STARCH DEBRANCHING ENZYMES

TECHNICAL FIELD

The present invention relates to enzymes having starch debranching activity. It further relates to nucleic acid encoding such enzymes, and methods of producing and using such enzymes and nucleic acid.

PRIOR ART

Starch is composed of highly branched (amylopectin), and lightly branched (amylose) glucan polymers arranged in a three-dimensional, semicrystalline structure, the starch granule. The degree of branching of amylopectin and the spatial organization of branches within the starch granule are very important in determining the physical properties of the starch and hence its value as a raw material for industry. The traditional view is that the branching pattern of amylopectin, and hence the way in which it is organised to form a granule, is determined by starch-branching enzymes which cleave short glucans from the non-reducing ends of chains and join them to residues within the same or an adjacent chain via α(1-6) linkages to form branches. There is, however, increasing evidence that the branching pattern of amylopectin results from the combined actions of branching and debranching enzymes.

"Debranching enzymes" hydrolyse α(1-6) glucosidic linkages in glucans. In plants, two quite different types have been described:

The "pullulanase" (EC 3.2.1.41) type is widely distributed in starch-degrading organs and in the chloroplasts of leaves. It is capable of the hydrolysis of the α(1-6) linkages of pullulan, amylopectin and α-

limit dextrins, but usually cannot hydrolyse glycogen.

The second type of debranching enzyme, the "isoamylase" (EC 3.2.1.68) type, has been described only in potato tubers and maize endosperm, but this is probably because there is, at the moment, no specific assay for isoamylase activity in crude extracts (i.e. where other hydrolysing enzymes may be present). It can hydrolyse the α(1-6) linkages of amylopectin, glycogen and α-limit dextrins, but not pullulan.

Evidence that debranching enzymes may be involved in determining amylopectin structure comes from analysis of the *sugary (su 1)* mutant of maize (Pan and Nelson 1984, James et al. 1995), the *sugary* mutant of rice (Nakamura et al. 1996a) and the *STA 7* mutant of *Chlamydomonas* (Mouille et al. 1996). All three mutations reduce or eliminate synthesis of conventional starch and cause the accumulation of a highly-branched, water-soluble glucan known as phytoglycogen. This change is accompanied by a reduction in the activity of debranching enzymes. In both maize and rice endosperm the activity of the pullulanase type of debranching enzyme is decreased, and in *Chlamydomonas* the activity of a debranching enzyme of unknown type disappears. In general terms, therefore, these phenotypes suggest that debranching enzyme is involved in determining the structure of amylopectin. However, understanding of the mutant phenotypes is far from complete.

Before the priority date of the present application, the *su1* locus from maize had been shown to encode a polypeptide which is very similar in amino-acid sequence to the bacterial isoamylase type of debranching enzyme, and not to pullulanases (James et al. 1995). Note, though, that the 5' end of the sequence was not necessarily complete in this publication. No effect of

the mutation on isoamylase activity in the endosperm was reported. The way in which the mutation brings about a decrease in pullulanase activity, and the relationship between this decrease and the accumulation of phytoglycogen were also not known.

After the priority date of the present application, nearly full-length maize SU1 was expressed in E. coli and purified. The recombinant enzyme was classified as an isoamylase (Rahman et al, 1998 Plant Physiol 117: 425-435).

Neither the rice nor the *Chlamydomonas* mutations have been fully characterised. In the former case, it has been established that the gene at the *sugary* locus does not encode the pullulanase that decreases in activity in the mutant endosperm (Nakamura et al. 1996b). In the latter case, the nature of the gene at the *STA7* locus is not known.

The general effects of these mutations form the basis for a new model to explain the synthesis of amylopectin and its organisation to form a granule (Ball et al. 1996). Briefly, it is proposed that debranching enzyme acts to "trim" a highly-branched phytoglycogen-like structure synthesised at the periphery of the growing granule. This creates the branching pattern typical of amylopectin which, unlike the branching pattern of phytoglycogen, allows the polymer to pack in an organised manner to form the semi-crystalline matrix of the granule.

A critical assessment of the validity of this model is not yet possible, in part because of the lack of understanding of the mutations on which it is based, and in part because of the lack of information about debranching enzymes generally, and in starch-synthesising organs in particular. The nature, number and

intracellular location of proteins with debranching activity is not known for any starch-synthesising organ, and sequences have been reported for only one plant isoamylase (the *su1* gene product) and a very few pullulanases. It is not known whether either isoamylase or pullulanase actually have the properties and specificities required by the Ball model.

Regardless of the validity of the Ball model, it seems highly likely that debranching enzymes play an important role in determining amylopectin structure, and hence in determining the physical properties of starch. The fact that the *su1* gene encodes an isoamylase suggests that this type of enzyme in particular may be involved. The decrease in pullulanase activity in the *su1* and *sugary* mutants also implicates this type of enzyme, and it has been reported (J. Kossmann and colleagues, MPI-MPP, Golm, Germany; verbal reports at open meetings) that modification of pullulanase activity in potato tubers brings about changes in the physical properties of the tuber starch.

Patent application WO 95/04826 [Kossmann et al] relates to a debranching enzyme obtained from potato. From the purification procedure used to obtain the amino acid sequence information it would appear that this relates to a single enzyme of the pullulanase type.

Patent application WO 95/03513 [Barry et al] relates to an isoamylase obtained from flavobacterium spp. The application does not disclose any corresponding enzymes or sequences from plants.

It can thus be seen that novel starch debranching enzymes, particularly those from plants, and particularly isoamylases, may provide a useful contribution to the art.

DISCLOSURE OF THE INVENTION

In a first aspect of the invention there is disclosed an isolated nucleic acid which comprises a nucleotide sequence which encodes a polypeptide which has the properties of an isoamylase, and is obtainable from Solanum tuberosum.

Preferably the nucleic acid molecule has the sequence shown in any of Seq ID Nos 1 to 3 or is degeneratively equivalent or complementary thereto.

Seq ID Nos 1 to 3 (Figs 1 to 3) represent nucleotide sequences derived by the present inventors from cDNA clones (designated 21, 15 and 9 respectively) from potato tubers and minitubers. Clone 15 came from a minituber library; clone 9 from a tuber library and clone 21 was found in both types of library. Each of these clones encodes all or part of an independent novel starch debranching enzyme.

The amino acid sequences for clones 21, 15 and 9 are given as Seq ID Nos 4-6 (Figs 4-6) respectively

The original nucleotide sequences for clones 21, 15 and 9 which were determined initially by the inventors are given as Seq ID Nos 10-12 (Figs 10-12) respectively. Owing to very minor variations in the sequencing process these differ at a very few positions from the sequences above: however in the case of clones 21 and 15 there is in excess of 99.5% identity between new and old sequences. Clone 9 has also been extended at its 3' terminus (still in excess of 99% identity). Corresponding amino acid sequences are at Seq ID Nos 13-15 (Figs 13-15) respectively.

Table 1

	Similarity	Identity
sul		
C9	63	46
C15	82	71
C21	58	35
Isopsean		
C9	53	32
C15	54	31
C21	48	23
Klepn		
C9	46.5	21.6
C15	48.5	23.2
C21	50.7	22.3
Kleae		
C9	47.3	21.6
C15	45.6	21.9
C21	49.1	21.8
Sopulspo		
C9	49.4	26.8
C15	43.6	22.1
C21	46.6	21.8
Puli		
C9	49.1	27.3
C15	50	26.6
C21	49.1	22.6

All of these sequences are unique, but show significant similarity at the level of predicted amino-acid sequence to the *sul* gene product of maize and the isoamylases of

micro-organisms.

A comparison of the clones was made with isoamylases (Table 1 - above); namely the *sul* gene (Sulzmay - EMBL Ac No U18908) and Isopseam (EMBL Ac No J03871; M28370) which is bacterial. Also with Pula_klepn (EMBL Ac No X52181; M32702) and Pula_kleae (EMBL Ac No M16187) which are bacterial pullulanases. Plus Sopulspo (EMBL Ac No X83969) which is a pullulanase from spinach.

Conserved domains I to IV occur in the amino acid sequences of clones 21, 15 and 9 at the following positions (Table 2):

Table 2

	Domain			
	I	II	III	IV
C9	352-357	426-434	467-470	535-540
C15	344-349	415-423	475-478	543-548
C21	455-460	515-523	556-559	623-628

This domain structure is typical of isoamylases; domain III does not occur in pullulanases.

The relationship between the various sequences is best illustrated by means of the dendogram (Fig 9) which shows the debranching enzymes in 2 groups, with clones 15, 9 and 21 all aligning with the isoamylases. Clone 15 is most similar to the *Sul* gene.

Since the clones show greater similarity to the isoamylase sequences, they have been putatively identified as isoamylases, and they are described as such

hereinafter; however it will be understood by the skilled person that the essence of the present invention is the making available of novel starch debranching enzymes, for instance for some of the purposes listed below, and this contribution to the art would not be diminished should the enzymes have properties not wholly consistent with the isoamylases described in the prior art. Indeed different properties and/or specificities may be advantageous for certain applications.

Thus the present inventors have for the first time demonstrated the existence of multiple forms of isoamylase in the potato tuber.

The nucleic acid molecules or vectors (see below) according to the present invention may be provided isolated and/or purified from their natural environment, in substantially pure or homogeneous form, or free or substantially free of nucleic acid or genes of the species of interest or origin other than the sequence encoding a polypeptide with the required function. Nucleic acid according to the present invention may include cDNA, RNA, genomic DNA and may be wholly or partially synthetic. "Nucleic acid" and "nucleic acid molecule" have the same meaning.

The term "isolate" encompasses all these possibilities. Where a DNA sequence is specified, e.g. with reference to a figure, unless context requires otherwise the RNA equivalent, with U substituted for T where it occurs, is encompassed.

Thus nucleic acid according to the present invention may comprise the sequence or complement of any one of Seq ID Nos. 1 to 3, including coding and/or non-coding regions where appropriate.

The disclosure of these sequences opens up for the first time the ability to manipulate the starch debranching activity in plants in a number of important respects. These include, *inter alia*, the ability to:

a) Reduce the activity of each of the isoamylases in potato tuber and other plants in which homologous enzymes are expressed.

b) Increase debranching enzyme activity in the potato tuber, by high level expression of one or more of each of the complete or partial potato cDNAs or sequences based thereon.

c) Alter the activity of each of the isoamylases in various different subcellular compartments (e.g. plastids or cytosol) or at various different developmental stages.

d) Study the effect of transformation experiments on the activities of isoforms of the debranching enzymes and related enzymes of starch synthesis and degradation, on the rates of starch synthesis, on starch structure, on the accumulation of soluble carbohydrates, and on the degradation of starch during sprouting.

e) Produce novel starch types in transgenic lines.

f) Produce novel isoamylases having modified activity.

g) Isolate corresponding isoamylases.

In essence the various nucleic acid molecules of the present invention may ultimately be used to promote or alter (in respect of the reaction catalysed) the nature of the starch debranching activity in a particular cell or organism. In some embodiments they may be used to repress starch debranching activity compared with that expressed in the untransformed cell or organism e.g. delay, retard, inhibit or slow down such activity.

In particular, alteration of debranching enzyme activity starch-synthesising cells would modify the structure of the starch accumulated in those cells in novel ways. The

modifications to the starch which may be achieved using the nucleic acid molecules of the present invention include:

1. Decrease the degree of branching of amylopectin thereby creating a starch that may swell less or form a stronger gel on heating in water. This may be achieved by increasing the activity of starch debranching enzymes, preferably through the over-expression of one, and most preferably more than one, endogenous or exogenous debranching enzymes.

2. Increase the degree of branching of starch thereby opening up the possibility of increasing its swelling properties and its ability to form a paste rather than a gel when heated in water. Particularly embraced is the production of phytoglycogen instead of some or all of the normal starch. This may be achieved by decreasing activity through the expression of antisense RNA.

3. Changing the branching pattern of amylopectin in other ways, thereby altering the physical properties of the starch. This may be accomplished by changing the isoform composition of debranching enzymes in a given tissue. Thus it may be achieved by selective decreases or increases in activity or subcellular localisation of endogenous isoamylases or the introduction of novel isoamylases and/or mutants, variants, derivatives or alleles thereof.

These and other aspects of the present invention will now be described in more detail.

Thus in a second aspect of the present invention there is disclosed a nucleic acid molecule encoding a mutant, variant, derivative or allele of a molecule of the first aspect, preferably to Seq ID Nos 1 to 3. Preferred

mutants, variants, derivatives and alleles are those which are homologous to the respective Seq ID No and which also encode a product which has the ability to promote starch debranching activity. Mutants, variants or derivatives of the complement of Seq ID Nos 1 to 3 are those which have the ability to repress starch debranching activity.

Methods for producing or identifying such a mutant, variant, derivative or allele (or other homologue) and assessing homology and function will now be discussed.

Changes to a sequence, to produce a mutant, variant or derivative, may be by way of one or more of addition, insertion, deletion or substitution of one or more nucleotides in the nucleic acid, leading to the addition, insertion, deletion or substitution of one or more amino acids in the encoded polypeptide. Of course, changes to the nucleic acid which make no difference to the encoded amino acid sequence (i.e. 'degeneratively equivalent') are included.

As is well-understood, homology at the amino acid level (i.e the encoded product of the nucleic acid molecule when expressed properly in frame) is generally in terms of amino acid similarity or identity. Similarity allows for "conservative variation", i.e. substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as arginine for lysine, glutamic for aspartic acid, or glutamine for asparagine. Similarity may be as defined and determined by the TBLASTN program, of Altschul *et al.* (1990) *J. Mol. Biol.* 215: 403-10, which is in standard use in the art, or, and this may be preferred, the standard program BestFit, which is part of the Wisconsin Package, Version 8, September 1994, (Genetics Computer Group, 575 Science

Drive, Madison, Wisconsin, USA, Wisconsin 53711). BestFit makes an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximize the number of matches using the local homology algorithm of Smith and Waterman

As is well known to those skilled in the art, altering the primary structure of a peptide by a conservative substitution may not significantly alter the activity of that peptide because the side-chain of the amino acid which is inserted into the sequence may be able to form similar bonds and contacts as the side chain of the amino acid which has been substituted out. This is so even when the substitution is in a region which is critical in determining the peptides conformation. Indeed, such changes may confer slightly advantageous properties on the peptide.

Also included are nucleic acids having a few "non-conservative" substitutions. As is well known to those skilled in the art, substitutions to regions of a peptide which are not critical in determining its conformation may not greatly affect its activity because they do not greatly alter the peptide's three dimensional structure. Those in important regions (e.g. conserved regions I to IV) may confer advantageous properties on the polypeptide product. Similarly it may be desirable to alter or otherwise manipulate the transit peptide sequence e.g. in clones 21 and 15, in order to alter the targeting or localisation properties of the enzymes.

A mutant, variant or derivative amino acid sequence in accordance with the present invention may include within the amino acid sequence encoded by Seq ID Nos 1 to 3 (see Figs 4 to 6) a single amino acid change with respect to the sequence shown or 2, 3, 4, 5, 6, 7, 8, or 9 changes, about 10, 15, 20, 30, 40 or 50 changes, or greater than

about 50, 60, 70, 80 or 90 changes.

In addition to one or more changes within the coding sequences of any one of Seq ID Nos 1 to 3, a mutant, variant or derivative nucleic acid molecule may have additional nucleotides at the 5' or 3' terminii. In particular it may be desirable to have a full length clone e.g. including any coding or non-coding regions (e.g. promoter) not included in the sequences but present in nature. These regions can be identified using methods analogous to those used to clone homologues or alleles as set out below.

In a third aspect of the present invention there is provided a method of identifying, mapping and/or cloning homologues or alleles from a plant species (including potato) which method employs all or part of the nucleotide sequence of Seq ID Nos 1 to 3. Suitable methods based on the sequences provided by the present invention are discussed below. If a portion of this a sequence is used this will be of sufficient length to identify homologues or alleles as described below.

Optionally, if a portion of nucleotide sequence is used, then this portion will not itself be identical to any part of Seq ID No 7 (Fig 7) which was used to detect clones 21, 15 and 9. Such a probe may therefore detect homologues and/or alleles which would not be detected using that Seq ID No 7.

In one embodiment of the third aspect, the nucleotide sequence of any one of Seq ID Nos 1 to 3, or any part thereof, may be used in a data-base search to find homologous sequences, expression products of which can be tested for ability to influence starch debranching, particularly for isoamylase activity. This may be achieved, for instance, using the vectors of the present

invention discussed hereinafter.

In a further embodiment of the third aspect, an isoamylase allele or homologue in accordance with the present invention is also obtainable by means of a method which includes providing a preparation of nucleic acid, e.g. from cells from a starch accumulating organ or tissue of a plant, providing a nucleic acid molecule having a nucleotide sequence shown in or complementary to a nucleotide sequence shown in any one of Seq ID Nos 1 to 3, preferably from within the coding sequence, contacting nucleic acid in said preparation with said nucleic acid molecule under conditions for hybridisation of said nucleic acid molecule to any said gene or homologue in said preparation, and identifying said gene or homologue if present by its hybridisation with said nucleic acid molecule.

Thus probing may employ the standard Southern blotting technique. For instance DNA may be extracted from cells and digested with different restriction enzymes. Restriction fragments may then be separated by electrophoresis on an agarose gel, before denaturation and transfer to a nitrocellulose filter. Labelled probe may be hybridised to the DNA fragments on the filter and binding determined. DNA for probing may be prepared from RNA preparations from cells.

Test nucleic acid may be provided from a cell as genomic DNA, cDNA or RNA, or a mixture of any of these, preferably as a library in a suitable vector. The information derived using genomic DNA may also be used in mapping, and in identifying associated non-expressed elements e.g. promoters.

Binding of a probe to target nucleic acid (e.g. DNA) may be measured using any of a variety of techniques at the

disposal of those skilled in the art. For instance, probes may be radioactively, fluorescently or enzymatically labelled. Other methods not employing labelling of probe include examination of restriction fragment length polymorphisms, amplification using PCR (see below), RN'ase cleavage and allele specific oligonucleotide probing.

Preliminary experiments may be performed by hybridising under low stringency conditions. For probing, preferred conditions are those which are stringent enough for there to be a simple pattern with a small number of hybridisations identified as positive which can be investigated further. It is well known in the art to increase stringency of hybridisation gradually until only a few positive clones remain. Suitable conditions would be achieved when a large number of hybridising fragments were obtained while the background hybridisation was low. Using these conditions nucleic acid libraries, e.g. cDNA libraries representative of expressed sequences, may be searched.

Those skilled in the art are well able to employ suitable conditions of the desired stringency for selective hybridisation, taking into account factors such as oligonucleotide length and base composition, temperature and so on. For instance, screening may initially be carried out under conditions, which comprise a temperature of about 37°C or less, a formamide concentration of less than about 50%, and a moderate to low salt (e.g. Standard Saline Citrate ('SSC') = 0.15 M sodium chloride; 0.15 M sodium citrate; pH 7) concentration.

Alternatively, a temperature of about 50°C or less and a high salt (e.g. 'SSPE'= 0.180 mM sodium chloride; 9 mM disodium hydrogen phosphate; 9 mM sodium dihydrogen

phosphate; 1 mM sodium EDTA; pH 7.4). Preferably the screening is carried out at about 37°C, a formamide concentration of about 20%, and a salt concentration of about 5 X SSC, or a temperature of about 50°C and a salt concentration of about 2 X SSPE. These conditions will allow the identification of sequences which have a substantial degree of homology (similarity, identity) with the probe sequence, without requiring the perfect homology for the identification of a stable hybrid.

Suitable conditions include, e.g. for detection of sequences that are about 80-90% identical, hybridization overnight at 42°C in 0.25M Na_2HPO_4, pH 7.2, 6.5% SDS, 10% dextran sulfate and a final wash at 55°C in 0.1X SSC, 0.1% SDS. For detection of sequences that are greater than about 90% identical, suitable conditions include hybridization overnight at 65°C in 0.25M Na_2HPO_4, pH 7.2, 6.5% SDS, 10% dextran sulfate and a final wash at 60°C in 0.1X SSC, 0.1% SDS.

Hybridisation is generally followed by identification of successful hybrids and then isolation of nucleic acid which has hybridised, which may involve one or more steps of PCR (see below).

Thus one part of the present invention is a probe for use in this method.

In a further embodiment of this aspect of the present invention, hybridisation of a nucleic acid molecule to an allele or homologue may be determined or identified indirectly, e.g using a nucleic acid amplification reaction, particularly the polymerase chain reaction (PCR). PCR requires the use of two primers to specifically amplify target nucleic acid, so preferably two nucleic acid molecules with sequences characteristic of any one of Seq ID Nos 1 to 3 are employed. However, if

RACE is used (see below) only one such specific primer may be needed. Characteristic in this sense is preferably in the sense of distinguishing them from known probes or sequences e.g. those associated with the *sul* gene.

PCR techniques for the amplification of nucleic acid are described in US Patent No. 4,683,195 and Saiki *et al. Science* 239: 487-491 (1988). PCR includes steps of denaturation of template nucleic acid (if double-stranded), annealing of primer to target, and polymerisation. The nucleic acid probed or used as template in the amplification reaction may be genomic DNA, cDNA or RNA. PCR may be used to amplify specific sequences from genomic DNA, specific RNA sequences and cDNA transcribed from mRNA. References for the general use of PCR techniques include Mullis et al, Cold Spring Harbor Symp. Quant. Biol., 51:263, (1987), Ehrlich (ed), PCR technology, Stockton Press, NY, 1989, Ehrlich et al, Science, 252:1643-1650, (1991), "PCR protocols; A Guide to Methods and Applications", Eds. Innis et al, Academic Press, New York, (1990).

Prior to any PCR that is to be performed, the complexity of a nucleic acid sample may be reduced where appropriate by creating a cDNA library for example using RT-PCR or by using the phenol emulsion reassociation technique (Clarke et al. (1992) NAR 20, 1289-1292) on a genomic library.

Thus a method involving use of PCR in obtaining nucleic acid according to the present invention may include providing a preparation of plant nucleic acid, providing a pair of nucleic acid molecule primers useful in (i.e. suitable for) PCR, at least one of said primers having a sequence shown in or complementary to all or part of a sequence shown in any one of Seq ID NOs 1 to 3, contacting nucleic acid in said preparation with said primers under conditions for performance of PCR,

performing PCR and determining the presence or absence of an amplified PCR product. The presence of an amplified PCR product may indicate identification of a gene of interest or fragment thereof.

Thus the methods of the invention may include hybridisation of one or more (e.g. two) probes or primers to target nucleic acid. Where the nucleic acid is double-stranded DNA, hybridisation will generally be preceded by denaturation to produce single-stranded DNA. The hybridisation may be as part of a PCR procedure, or as part of a probing procedure not involving PCR. An example procedure would be a combination of PCR and low stringency hybridisation. A screening procedure, chosen from the many available to those skilled in the art, is used to identify successful hybridisation events and isolated hybridised nucleic acid.

The primers for use in these methods form one part of the present invention.

In any case, an oligonucleotide for use in probing or nucleic acid amplification may have about 10 or fewer codons (e.g. 6, 7 or 8), i.e. be about 30 or fewer nucleotides in length (e.g. 18, 21 or 24). Generally specific primers are upwards of 14 nucleotides in length. For optimum specificity, primers of 16-24 nucleotides in length may be preferred. Those skilled in the art are well versed in the design of primers for use processes such as PCR.

In all cases the nucleic acids of the second aspect, or identified using the third aspect, share homology with those of the first aspect. Homology may be at the nucleotide sequence and/or amino acid sequence level. Preferably, there is at least about 83% homology, most preferably at least about 85%, 90%, 95%, 96%, 97%, 98%,

99% or 99.5% homology.

Homology may be over the full-length of the relevant sequence shown herein, or may more preferably be over a contiguous sequence of about or greater than about 20, 25, 30, 33, 40, 50, 67, 133 or more amino acids or codons, compared with the relevant amino acid sequence or nucleotide sequence as the case may be.

Similarly the mutant, variant, derivative or allele (or other homologue) in accordance with the present invention will promote, alter or repress the starch debranching activity of a cell into which it is introduced.

One possible mode of analysis of this activity is by transformation to assess function on introduction into a plant, plant cell or other cell of interest; methodology for such transformation is described in more detail below.

The nucleic acid of the present invention, which may contain for example DNA corresponding to any one of Seq ID Nos 1 to 3, may be in the form of a recombinant and preferably replicable vector.

Such vectors form a fourth aspect of the present invention.

DNA vector is defined to include, inter alia, any plasmid, cosmid, phage or *Agrobacterium* binary vector in double or single stranded linear or circular form which may or may not be self transmissible or mobilizable. Can transform prokaryotic or eukaryotic host either by integration into the cellular genome or exist extrachromosomally (e.g. autonomous replicating plasmid with an origin of replication). Vectors may be introduced into hosts by any appropriate method e.g. conjugation,

mobilisation, transformation, transfection, transduction or electoporation. Also included are shuttle vectors by which is meant a DNA vehicle capable, naturally or by design, of replication in both the actinomycetes and related species and in bacteria and/or eucaryotic (e.g. higher plant, mammalian, yeast or fungal cells).

A vector including nucleic acid according to the present invention need not include a promoter or other regulatory sequence, particularly if the vector is to be used to introduce the nucleic acid into cells for recombination into the genome.

However, in a preferred embodiment of the fourth aspect the vector is an expression vector. Those skilled in the art are well able to construct vectors and design protocols for recombinant gene expression. Suitable vectors can be chosen or constructed, containing appropriate regulatory sequences, including promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other sequences as appropriate. For further details see, for example, *Molecular Cloning: a Laboratory Manual:* 2nd edition, Sambrook *et al*, 1989, Cold Spring Harbor Laboratory Press. Many known techniques and protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into cells and gene expression, and analysis of proteins, are described in detail in *Current Protocols in Molecular Biology*, Second Edition, Ausubel et al. eds., John Wiley & Sons, 1992. The disclosures of Sambrook et al. and Ausubel et al. are incorporated herein by reference. Specific procedures and vectors previously used with wide success upon plants are described by Bevan (Nucl. Acids Res. 12, 8711-8721 (1984)) and Guerineau and Mullineaux (1993) (Plant transformation and expression vectors. In: Plant

Molecular Biology Labfax (Croy RRD ed) Oxford, BIOS Scientific Publishers, pp 121-148).

Selectable genetic markers may be used consisting of chimaeric genes that confer selectable phenotypes such as resistance to antibiotics such as kanamycin, hygromycin, phosphinotricin, chlorsulfuron, methotrexate, gentamycin, spectinomycin, imidazolinones and glyphosate.

Thus nucleic acid molecules of the present invention may be under the control of an appropriate promoter or other regulatory elements for expression in a host cell such as a microbial, e.g. bacterial, or plant cell. In the case of genomic DNA, this may contain its own promoter or other regulatory elements and in the case of cDNA this may be under the control of an appropriate promoter or other regulatory elements for expression in the host cell.

Thus one aspect the present invention provides a gene construct, preferably a replicable vector, comprising a promoter operatively linked to a nucleotide sequence provided by the present invention, e.g. any one of Seq ID Nos 1 to 3, the complement, or any mutant, variant or allele thereof.

By "promoter" is meant a sequence of nucleotides from which transcription may be initiated of DNA operably linked downstream (i.e. generally in the 3' direction on the sense strand of double-stranded DNA).

"Operably linked" means joined as part of the same nucleic acid molecule, suitably positioned and oriented for transcription to be initiated from the promoter. DNA operably linked to a promoter is "under transcriptional initiation regulation" of the promoter.

The promoter may include one or more sequence motifs or elements conferring developmental and/or tissue-specific regulatory control of expression. Other regulatory sequences may be included, for instance as identified by mutation or digest assay in an appropriate expression system or by sequence comparison with available information, e.g. using a computer to search on-line databases. Sequences for intra- or intercellular targetting may also be included e.g. plastid targetting sequences as described in, or modified from, Stark et al (1992) Science 258: 287-292. Also included may be appropriate untranscribed regions e.g. which cause the addition of the polyadenylate nucleotides to 3' end of transcribed RNA.

Suitable promoters may include the Cauliflower Mosaic Virus 35S (CaMV 35S) gene promoter that is expressed at a high level in virtually all plant tissues (Benfey et al, 1990a and 1990b). Other promoters may include the tuber specific B33 promoter (Rocha-Sosa et al (1989) EMBO J 8:23-29), or the patatin (class I) promoter.

In one embodiment of the fourth aspect these is disclosed a gene construct, preferably a replicable vector, comprising an inducible promoter operatively linked to a nucleotide sequence provided by the present invention.

The present invention also provides plants transformed with said gene construct and methods comprising introduction of such a construct into a plant cell and/or induction of expression of a construct within a plant cell, by application of a suitable stimulus, an effective exogenous inducer.

The term "inducible" as applied to a promoter is well understood by those skilled in the art. In essence, expression under the control of an inducible promoter is

"switched on" or increased in response to an applied stimulus. The nature of the stimulus varies between promoters. Some inducible promoters cause little or undetectable levels of expression (or no expression) in the absence of the appropriate stimulus. Other inducible promoters cause detectable constitutive expression in the absence of the stimulus. Whatever the level of expression is in the absence of the stimulus, expression from any inducible promoter is increased in the presence of the correct stimulus. The preferable situation is where the level of expression increases upon application of the relevant stimulus by an amount effective to alter a phenotypic characteristic. Thus an inducible (or "switchable") promoter may be used which causes a basic level of expression in the absence of the stimulus which level is too low to bring about a desired phenotype (and may in fact be zero). Upon application of the stimulus, expression is increased (or switched on) to a level which brings about the desired phenotype.

A suitable inducible promoter is the GST-II-27 gene promoter which has been shown to be induced by certain chemical compounds which can be applied to growing plants. The promoter is functional in both monocotyledons and dicotyledons. It can therefore be used to control gene expression in a variety of genetically modified plants, including field crops such as canola, sunflower, tobacco, sugarbeet, cotton; cereals such as wheat, barley, rice, maize, sorghum; fruit such as tomatoes, mangoes, peaches, apples, pears, strawberries, bananas, and melons; and vegetables such as carrot, lettuce, cabbage and onion. The GST-II-27 promoter is also suitable for use in a variety of tissues, including roots, leaves, stems and reproductive tissues.

The vectors of the fourth aspect of the invention may be

used, *inter alia*, to transform plants and plant cells thereby altering their properties in a number of important respects.

Thus specific changes in activity of individual forms of isoamylase may be brought about by sense or antisense transformation. The disclosure by the present inventors of several forms of isoamylase has opened up the possibility of 'fine tuning' their effect(s) on amylopectin structure and hence on the properties of starch, depending upon which isoform is changed, the degree to which the activity is increased or decreased, and the timing of this change in relation to the period of starch synthesis in the tuber. Changes in activity of more than one isoform simultaneously can likewise be used to produce unique effects on amylopectin structure and hence on the properties of starch. Similarly the introduction via transformation of one or more of the isoamylases from the potato tuber into starch-synthesising organs of <u>other</u> species may be used to bring about unique and novel changes in the structure of amylopectin and hence in the properties of starch in those organs.

Thus in a fifth aspect of the present invention there is disclosed a host cell containing nucleic acid or a vector according to the present invention, especially a plant or a microbial cell.

This aspect of the present invention further encompasses a host cell transformed with nucleic acid or a vector according to the present invention, especially a plant or a microbial cell. Within the cell, the nucleic acid may be incorporated within the chromosome. There may be more than one such heterologous nucleotide sequence per haploid genome.

When introducing a chosen gene construct into a cell, certain considerations must be taken into account, well known to those skilled in the art. The nucleic acid to be inserted should be assembled within a construct which contains effective regulatory elements which will drive transcription. There must be available a method of transporting the construct into the cell. Once the construct is within the cell membrane, integration into the endogenous chromosomal material either will or will not occur. Finally, as far as plants are concerned the target cell type must be such that cells can be regenerated into whole plants (see below).

Plants transformed with the DNA segment containing the sequence may be produced by standard techniques which are already known for the genetic manipulation of plants. DNA can be transformed into plant cells using any suitable technology, such as a disarmed Ti-plasmid vector carried by Agrobacterium exploiting its natural gene transfer ability (EP-A-270355, EP-A-0116718, NAR 12(22) 8711 - 87215 1984), particle or microprojectile bombardment (US 5100792, EP-A-444882, EP-A-434616) microinjection (WO 92/09696, WO 94/00583, EP 331083, EP 175966, Green *et al.* (1987) *Plant Tissue and Cell Culture*, Academic Press), electroporation (EP 290395, WO 8706614 Gelvin Debeyser - see attached) other forms of direct DNA uptake (DE 4005152, WO 9012096, US 4684611), liposome mediated DNA uptake (e.g. Freeman *et al. Plant Cell Physiol.* 29: 1353 (1984)), or the vortexing method (e.g. Kindle, *PNAS U.S.A.* 87: 1228 (1990d). Physical methods for the transformation of plant cells are reviewed in Oard, 1991, *Biotech. Adv.* 9: 1-11.

Agrobacterium transformation is widely used by those skilled in the art to transform dicotyledonous species. Recently, there has been substantial progress towards the routine production of stable, fertile transgenic plants

in almost all economically relevant monocot plants (Toriyama, et al. (1988) *Bio/Technology* 6, 1072-1074; Zhang, *et al.* (1988) *Plant Cell Rep.* 7, 379-384; Zhang, *et al.* (1988) *Theor Appl Genet* 76, 835-840; Shimamoto, *et al.* (1989) *Nature* 338, 274-276; Datta, *et al.* (1990) *Bio/Technology* 8, 736-740; Christou, *et al.* (1991) *Bio/Technology* 9, 957-962; Peng, *et al.* (1991) International Rice Research Institute, Manila, Philippines 563-574; Cao, *et al.* (1992) *Plant Cell Rep.* 11, 585-591; Li, *et al.* (1993) *Plant Cell Rep.* 12, 250-255; Rathore, *et al.* (1993) *Plant Molecular Biology* 21, 871-884; Fromm, *et al.* (1990) *Bio/Technology* 8, 833-839; Gordon-Kamm, *et al.* (1990) *Plant Cell* 2, 603-618; D'Halluin, *et al.* (1992) *Plant Cell* 4, 1495-1505; Walters, *et al.* (1992) *Plant Molecular Biology* 18, 189-200; Koziel, *et al.* (1993) *Biotechnology* 11, 194-200; Vasil, I. K. (1994) *Plant Molecular Biology* 25, 925-937; Weeks, *et al.* (1993) *Plant Physiology* 102, 1077-1084; Somers, *et al.* (1992) *Bio/Technology* 10, 1589-1594; WO92/14828). In particular, *Agrobacterium* mediated transformation is now emerging also as an highly efficient alternative transformation method in monocots (Hiei *et al.* (1994) *The Plant Journal* 6, 271-282).

The generation of fertile transgenic plants has been achieved in the cereals rice, maize, wheat, oat, and barley (reviewed in Shimamoto, K. (1994) *Current Opinion in Biotechnology* 5, 158-162.; Vasil, *et al.* (1992) *Bio/Technology* 10, 667-674; Vain et al., 1995, *Biotechnology Advances* 13 (4): 653-671; Vasil, 1996, *Nature Biotechnology* 14 page 702).

Microprojectile bombardment, electroporation and direct DNA uptake are preferred where Agrobacterium is inefficient or ineffective. Alternatively, a combination of different techniques may be employed to enhance the efficiency of the transformation process, eg bombardment

with Agrobacterium coated microparticles (EP-A-486234) or microprojectile bombardment to induce wounding followed by co-cultivation with Agrobacterium (EP-A-486233).

Following transformation, a plant may be regenerated, e.g. from single cells, callus tissue or leaf discs, as is standard in the art. Almost any plant can be entirely regenerated from cells, tissues and organs of the plant. Available techniques are reviewed in Vasil et al., *Cell Culture and Somatic Cell Genetics of Plants, Vol I, II and III, Laboratory Procedures and Their Applications*, Academic Press, 1984, and Weissbach and Weissbach, *Methods for Plant Molecular Biology*, Academic Press, 1989.

The particular choice of a transformation technology will be determined by its efficiency to transform certain plant species as well as the experience and preference of the person practising the invention with a particular methodology of choice. It will be apparent to the skilled person that the particular choice of a transformation system to introduce nucleic acid into plant cells is not essential to or a limitation of the invention, nor is the choice of technique for plant regeneration.

Thus a sixth aspect of the present invention provides a method of generating a cell involving introduction of a vector as described in relation to the fourth aspect above into plant cell and causing or allowing recombination between the vector and the cell genome to introduce the sequence of nucleotides into the genome. Preferably the cell is a plant cell.

Thus according to the invention there is provided a plant cell having incorporated into its genome nucleic acid, particularly heterologous nucleic acid, as provided by the present invention, under operative control of a

regulatory sequence for control of expression. The coding sequence may be operably linked to one or more regulatory sequences which may be heterologous or foreign to the gene i.e. which is not naturally associated with the gene for its expression. The nucleic acid according to the invention may be placed under the control of an externally inducible gene promoter to place expression under the control of the user.

The term "heterologous" may be used to indicate that the gene/sequence of nucleotides in question have been introduced into said cells of the plant or an ancestor thereof, using genetic engineering, i.e. by human intervention. A transgenic plant cell, i.e. transgenic for the nucleic acid in question, may be provided. The transgene may be on an extra-genomic vector or incorporated, preferably stably, into the genome. A heterologous gene may replace an endogenous equivalent gene, ie one which normally performs the same or a similar function, or the inserted sequence may be additional to the endogenous gene or other sequence. An advantage of introduction of a heterologous gene is the ability to place expression of a sequence under the control of a promoter of choice, in order to be able to influence expression according to preference. Furthermore, mutants, variants and derivatives of the wild-type gene, e.g. with higher or lower activity than wild-type, may be used in place of the endogenous gene. Nucleic acid heterologous, or exogenous or foreign, to a plant cell may be non-naturally occurring in cells of that type, variety or species. Thus, nucleic acid may include a coding sequence of or derived from a particular type of plant cell or species or variety of plant, placed within the context of a plant cell of a different type or species or variety of plant. A further possibility is for a nucleic acid sequence to be placed within a cell in which it or a homologue is found naturally, but wherein

the nucleic acid sequence is linked and/or adjacent to nucleic acid which does not occur naturally within the cell, or cells of that type or species or variety of plant, such as operably linked to one or more regulatory sequences, such as a promoter sequence, for control of expression. A sequence within a plant or other host cell may be identifiably heterologous, exogenous or foreign.

A plant may be regenerated from one or more transformed plant cells. Thus a plant including a plant cell according to the invention forms a seventh aspect of the present invention, along with any part or propagule thereof, seed, selfed or hybrid progeny and descendants.

A plant according to the present invention may be one which does not breed true in one or more properties. Plant varieties may be excluded, particularly registrable plant varieties according to Plant Breeders' Rights. It is noted that a plant need not be considered a "plant variety" simply because it contains stably within its genome a transgene, introduced into a cell of the plant or an ancestor thereof.

Preferred plants of the present invention include modified potato, pea, maize, wheat, cassava, rice and barley.

In addition to a plant, the present invention provides any clone of such a plant, seed, selfed or hybrid progeny and descendants, and any part of any of these, such as cuttings, seed. The invention provides any plant propagule, that is any part which may be used in reproduction or propagation, sexual or asexual, including cuttings, seed and so on. Also encompassed by the invention is a plant which is a sexually or asexually propagated off-spring, clone or descendant of such a plant, or any part or propagule of said plant, off-

spring, clone or descendant.

As discussed above, particularly embraced by the present invention are methods of influencing or affecting the starch debranching activities of a plant comprising the use of any of the nucleic acids, vectors and/or other materials or methods discussed in relation to aspects one to seven above, including causing or allowing expression of a heterologous nucleic acid sequence within cells of the plant. Such methods form an eighth aspect of the present invention.

In one embodiment there is provided a method including expression of a nucleic acid molecule having a sequence identical or complementary to all or part of Seq ID Nos 1 to 3, or a mutant, variant, allele or other derivative of the sequence, within cells of a plant (thereby producing the encoded polypeptide), following an earlier step of introduction of the nucleic acid into a cell of the plant or an ancestor thereof. Such a method may be used to influence the starch generated within the cells of that plant.

In the present invention, over-expression may be achieved by introduction of the nucleic acid molecules discussed above in a sense orientation. Thus, the present invention provides a method of influencing the starch debranching activity of a plant, the method including causing or allowing expression of the product (polypeptide or nucleic acid transcript) encoded by heterologous nucleic acid according to the invention from that nucleic acid within cells of the plant.

Conversely, down-regulation of expression of a target gene (i.e. an isoamylase encoded by any of the nucleic acid molecules of the present invention) may be achieved using anti-sense technology or "sense regulation" ("co-

suppression").

In using anti-sense genes or partial gene sequences to down-regulate gene expression, a nucleotide sequence is placed under the control of a promoter in a "reverse orientation" such that transcription yields RNA which is complementary to normal mRNA transcribed from the "sense" strand of the target gene. See, for example, Rothstein et al, 1987; Smith *et al*, (1988) *Nature* 334, 724-726; Zhang *et al*, (1992) *The Plant Cell* 4, 1575-1588, English *et al*., (1996) *The Plant Cell* 8, 179-188. Antisense technology is also reviewed in Bourque, (1995), *Plant Science* 105, 125-149, and Flavell, (1994) *PNAS USA* 91, 3490-3496.

An alternative is to use a copy of all or part of the target gene inserted in sense, that is the same, orientation as the target gene, to achieve reduction in expression of the target gene by co-suppression.

Both of these methods will now be discussed in more detail.

The complete sequence corresponding to the coding sequence of the targeted isoamylase(in reverse orientation for anti-sense) need not be used. For example fragments of sufficient length may be used. It is a routine matter for the person skilled in the art to screen fragments of various sizes and from various parts of the coding sequence to optimise the level of anti-sense inhibition. It may be advantageous to include the initiating methionine ATG codon, and perhaps one or more nucleotides upstream of the initiating codon. A further possibility is to target a conserved sequence of a gene, e.g. a sequence that is characteristic of one or more genes, such as a regulatory sequence.

The sequence employed may be about 500 nucleotides. However it may be less, possibly about 400 nucleotides, about 300 nucleotides, about 200 nucleotides, or about 100 nucleotides. It may be even be possible to use oligonucleotides of much shorter lengths, 14-23 nucleotides. Longer fragments, for instance longer than about 500 nucleotides are preferable where possible, such as longer than about 600 nucleotides, than about 700 nucleotides, than about 800 nucleotides, than about 1000 nucleotides or more.

It may be preferable that there is complete sequence identity in the sequence used for down-regulation of expression of a target sequence, and the target sequence, though total complementarity or similarity of sequence is not essential. One or more nucleotides may differ in the sequence used from the target gene. Thus, a sequence employed in a down-regulation of gene expression in accordance with the present invention may be a wild-type sequence (e.g. gene) selected from those available, or a mutant, derivative, variant or allele, by way of insertion, addition, deletion or substitution of one or more nucleotides, of such a sequence. The sequence need not include an open reading frame or specify an RNA that would be translatable. It may be preferred for there to be sufficient homology for the respective anti-sense and sense RNA molecules to hybridise. There may be down regulation of gene expression even where there is about 5%, 10%, 15% or 20% or more mismatch between the sequence used and the target gene, although it may be advantageous to have minimal mismatch.

Thus generally speaking, the transcribed nucleic acid may represent a fragment of an isoamylase gene, such as any one of those corresponding to Seq ID Nos 1 to 3, or the complement thereof, or may be a mutant, derivative, variant or allele thereof, in similar terms as discussed

above in relation to alterations being made to an coding sequence and the homology of the altered sequence (see the first or second aspects of the invention). The homology may be sufficient for the transcribed anti-sense RNA to hybridise with nucleic acid within cells of the plant, though irrespective of whether hybridisation takes place the desired effect is down-regulation of gene expression.

Anti-sense regulation may itself be regulated by employing an inducible promoter in an appropriate construct. Thus, the present invention also provides a method of influencing a starch debranching activity of a plant, the method including causing or allowing anti-sense transcription from heterologous nucleic acid according to the invention within cells of the plant.

The anti-sense constructs (nucleic acids) themselves are also embraced by the present invention, as is use of these constructs for down-regulation of gene expression, particularly down-regulation of expression of an isoamylase or homologue thereof, preferably in order to influence the starch debranching enzyme activity of a plant, especially a crop plant.

As discussed above, when additional copies of the target gene are inserted in sense, that is the same, orientation as the target gene, a range of phenotypes is produced which includes individuals where over-expression occurs and some where under-expression of protein from the target gene occurs. When the inserted gene is only part of the endogenous gene the number of under-expressing individuals in the transgenic population increases. The mechanism by which sense regulation occurs, particularly down-regulation, is not well-understood. However, this technique is well-reported in scientific and patent literature and is used routinely for gene control. See,

for example, van der Krol *et al.*, (1990) *The Plant Cell 2*, 291-229; Napoli *et al.*, (1990) The Plant Cell *2*, 279-289; Zhang et al, 1992 *The Plant Cell 4*, 1575-1588, and US-A-5,231,020

Again, anti-sense fragments, mutants and so on may be used in similar terms as described above in relation to the second aspect.

Further methods of down-regulating activity include inhibition by expressing dominant negative versions (i.e. mutant or truncated versions) of the isoamylases which will inhibit endogenous, wild-type enzymes in a competitive or non-competitive way e.g. by competing for binding sites on the starch granules, or by association to form non-functional multimers. Alternatively one can use ribozymes, e.g. hammerhead ribozymes, which can catalyse the site-specific cleavage of RNA, such as mRNA (see e.g. Jaeger (1997) "The new world of ribozymes" Curr Opin Struct Biol 7:324-335, or Gibson & Shillitoe (1997) "Ribozymes: their functions and strategies form their use" Mol Biotechnol 7: 242-251.)

Thus, the present invention also provides a method of influencing a starch debranching activity of a plant, the method including causing or allowing transcription of nucleic acid as described above, within cells of the plant.

Here the starch debranching activity of the product is preferably suppressed as a result of under-expression of isoamylase within the plant cells.

In a ninth aspect of the invention there is disclosed the expression product (preferably being an isoamylase) of any of the nucleic acid sequences disclosed above, particularly those of the first and second aspects of the

invention, optionally by means of the vectors of the fourth aspect. Example amino acid sequences are given in Figs 4 to 6. Also embraced are methods of generating isoamylases by expression from encoding nucleic acid therefore under suitable conditions, which may be in suitable host cells. Following expression, the product may be isolated from the expression system and may be used as desired, for instance in formulation of a composition including at least one additional component.

One particular use for such expression products may be raising antibodies. Such antibodies form a tenth aspect of the present invention.

Thus purified protein of the ninth aspect, or a fragment, mutant, derivative or variant thereof, e.g. produced recombinantly by expression from encoding nucleic acid, may be used to raise antibodies employing techniques which are standard in the art. Antibodies and polypeptides comprising antigen-binding fragments of antibodies may be used in identifying homologues from other species as discussed further below.

Methods of producing antibodies include immunising a mammal (e.g. human, mouse, rat, rabbit, horse, goat, sheep or monkey) with the protein or a fragment thereof. Antibodies may be obtained from immunised animals using any of a variety of techniques known in the art, and might be screened, preferably using binding of antibody to antigen of interest. For instance, Western blotting techniques or immunoprecipitation may be used (Armitage et al, 1992, Nature 357: 80-82). Antibodies may be polyclonal or monoclonal.

As an alternative or supplement to immunising a mammal, antibodies with appropriate binding specificity may be obtained from a recombinantly produced library of

expressed immunoglobulin variable domains, e.g. using lambda bacteriophage or filamentous bacteriophage which display functional immunoglobulin binding domains on their surfaces; for instance see WO92/01047.

Antibodies raised to a polypeptide or peptide can be used, *inter alia*, in the identification and/or isolation and/or localisation (e.g. intracellular) of the peptides of the present invention and homologous polypeptides, and may also permit isolation of the encoding genes. Thus, the present invention provides a method of identifying or isolating a polypeptide with starch debranching function (in accordance with embodiments disclosed herein), comprising screening candidate polypeptides with a polypeptide comprising the antigen-binding domain of an antibody (for example whole antibody or a fragment thereof) which is able to bind an starch debranching polypeptide or fragment, variant or derivative thereof or preferably has binding specificity for such a polypeptide. Specific binding members such as antibodies and polypeptides comprising antigen binding domains of antibodies that bind and are preferably specific for an isoamylase or mutant, variant or derivative thereof represent further aspects of the present invention, as do their use and methods which employ them.

Candidate polypeptides for screening may for instance be the products of an expression library created using nucleic acid derived from an plant of interest, or may be the product of a purification process from a natural source. A polypeptide found to bind the antibody may be isolated and then may be subject to amino acid sequencing. Any suitable technique may be used to sequence the polypeptide either wholly or partially (for instance a fragment of the polypeptide may be sequenced). Amino acid sequence information may be used in obtaining nucleic acid encoding the polypeptide, for instance by

designing one or more oligonucleotides (e.g. a degenerate pool of oligonucleotides) for use as probes or primers in hybridization to candidate nucleic acid, or by searching computer sequence databases, as discussed above.

An eleventh aspect of the present invention is a polysaccharide generated (*in vivo* or using an enzyme preparation) by a process comprising the use of an isoamylase of the ninth aspect. Also embraced is starch produced in the transformed plants and cells discussed above. Such starch is preferably derived from amylopectin but has any of a decreased, increased or otherwise altered degree of branching, with a corresponding alteration in properties e.g. swelling or ability to form a paste rather than a gel when heated in water. Commodities (e.g. foodstuffs) comprising such starches form a further aspect of the present invention.

Other commodities which may benefit from the modified starches of the present invention include biodegradable plastics; food-processing thickeners; starch coated films, papers & textiles; paint thickeners; mining explosives; pharmaceuticals and glues. The modified starches can be used analagously to prior art starches in these materials, in ways which are well known to those skilled in the respective technical fields.

The invention will now be further illustrated with reference to the following non-limiting Figures and Examples. Other embodiments falling within the scope of the invention will occur to those skilled in the art in the light of these.

FIGURES

Fig 1a and 1b shows nucleotide Seq ID No 1 (from isoamylase clone 21).

Fig 2a and 2b shows nucleotide Seq ID No 2 (from isoamylase clone 15).

Fig 3a and 3b shows nucleotide Seq ID No 3 (from isomylase clone 9).

Fig 4 shows amino acid Seq ID No 4 (from isoamylase clone 21).

Fig 5 shows amino acid Seq ID No 5 (from isoamylase clone 15).

Fig 6 shows amino acid Seq ID No 6 (from isomylase clone 9).

Fig 7 shows Seq ID No 7, corresponding to the Arabidopsis thaliana v. columbia probe At69012.new_est taken from the Medline Database, originally published by Newman et al (1994) Plant Physiol 106: 1241-1255.

Fig 8(a) shows the transit peptide sequence from clone 21 (Seq ID No 8). Fig 8(b) shows the transit peptide sequence from clone 15 (Seq ID No 9).

Fig 9 shows a dendogram which places the debranching enzymes in 2 distinct groups, with clones 15, 9 and 21 all aligning with the isoamylases.

Fig 10a and 10b shows nucleotide Seq ID No 10 (original sequence from isoamylase clone 21).

Fig 11a and 11b shows nucleotide Seq ID No 11 (original sequence from isoamylase clone 15).

Fig 12 shows nucleotide Seq ID No 12 (original sequence from isomylase clone 9).

Fig 13 shows amino acid Seq ID No 13 (original sequence from isoamylase clone 21).

Fig 14 shows amino acid Seq ID No 14 (original sequence from isoamylase clone 15). 'x' is unknown aminoacid. '*' is a stop codon.

Fig 15 shows amino acid Seq ID No 15 (original sequence from isomylase clone 9).

Fig 16(a) shows the forward primer (Seq ID No 16) used to isolate the probe (see Examples below). Fig 16(b) shows the reverse primer (Seq ID No 17).

EXAMPLES

Example 1- cloning of the debranching enzymes from potato

Briefly, cDNA clones from potato were isolated from cDNA libraries synthesised from mRNA from both developing tubers and from *in vitro* grown minitubers. The probe was an EST from *Arabidopsis* (At69012.new_est) which was identified by the present inventors as showing significant homology to the *sul* gene from maize.

Libraries:
Two independent libraries were prepared. These were -

a) from mRNA from developing tubers from greenhouse cultivated potato plants (Solanum tuberosum var desiree)

b) from mRNA from minitubers induced on stem explants of potato(Solanum tuberosum var desiree) cultured in vitro according to the method of Visser et al (1994) Physiol Plantarum 90: 285-292. Minitubers were used in addition to tubers in order to assess starch synthesising organ which has a different gene expression profile to tubers.

The cDNA was synthesized by reverse transcription using as a template poly(A)RNA passed twice over oligo(dT)cellulose. The poly(A)RNA was reverse transcribed to form the first strand cDNA and the second strand was prepared using DNA polymerase I (large subunit), T4 DNA polymerase and RNAase H as described in the manufacturer's instructions for the cDNA synthesis kit (Amersham plc, UK).

cDNA was ligated to EcoRI adaptors as described in the Rapid adapter ligation kit (Amersham plc, UK) and then cloned into the EcoRI of the λ cloning vector λgt10 according to the manufacturers instructions for the λgt10 cloning kit (Amersham plc, UK).

Probe:

The probe used was a fragment of an *Arabidopsis* EST (EMBL ID No. At69012.new_est; accession no. H36690). This EST was identified initially using a BLAST search of EST databases.

In order to determine the extent of the homology between the *Arabidopsis* EST and the *Su1* gene product, the EST was further sequenced using an Applied Biosystems Taq cycle sequencing kit (Perkin Elmer) and an ABI automated sequencer. This generated a further 780 bp sequence data. It showed significant homology to the deduced amino acid sequence for the product of the *Su1* gene from maize that encodes an isoamylase type of debranching enzyme.

The probe was prepared using PCR amplification of a miniprep of the plasmid. The PCR used the M13 reverse primer:

5'-CAG GAA ACA GCT ATG AC -3' (SEq ID No 5)

And also a primer specific for the 3' end of the EST, at a point before the polyA tail. This was designated G3712:

5'-GAT CAT AAC TTG AGT TCT AAG CGG -3'

The amplified fragment was cut with PstI to remove the sequences from the vector. The fragment was purified and then labelled using an oligonucleotide random priming labelling kit to provide the probe.

Screening:

Approximately 60,000 plaques from the tuber library (unamplified) and 60,000 plaques from the minituber library (unamplified) were used to infect E. coli (strain NM514) and the resultant plaques were screened using a 1.2 kb fragment of EST cDNA clone (At69012.new_est) which lacked the poly(A) tail. Filters were subsequently washed at low stringency (2xSSC, 0.5%SDS, 55°C for two washes).

16 independent phages (5 from tuber and 11 from minituber) that showed different levels of hybridization to the EST probe were selected.

Subcloning:

DNA from 9 independent clones was subcloned into either pCR2.1 (Invitrogen) or pBluescript (Stratagene) in *E. coli*. Those in pCR2.1 were subcloned following PCR amplification of the inserts using λgt10 specific oligonucleotides. Those in pBluescript were isolated as EcoRI fragments from λDNA preparations.

Sequencing & analysis:

Clones were sequenced using the Taq cycle sequencing kit from Perkin Elmer and the ABI automated sequencer. To complete any incomplete sequence, primers based on the

known portions of the sequence are used to 'walk' along the clones in the library to identify the remaining portions. Following initial sequencing of the C9 clone, a longer cDNA was obtained and sequenced.

The predicted N-terminal amino acid sequences for C15 and C21 fit the criteria for plastid transit peptides. A summary of the cDNA clones is presented below. This refers to the original sequences. Corresponding comparisons with *sul* for the new sequences are shown above.

Table 3

cDNA clones	length (kb)	open reading frame (number of amino acids)		relationship to sul (%, at amino acid level)	
		total	predicted transit peptide	similarity	identity
C9	2.6	766	none	61	45
C15	2.7	793	47	82	70
C21	2.9	878	38	57	35

Example 2 - Transformation and antisense constructs

The clones encoding the isoamylases are used to construct a series of lines of antisense potato plants. The clone (C9, C15 and C21) is subcloned in antisense orientation between the CaMV 35S promoter and the CaMV terminator sequences of pJIT60 (see Guerineau & Mullinieaux (1993)in Plant Molecular Biology Lab Fax ed. Croy RRD BIOS Scientific, Oxford, UK pp 121-148). This construct has been subcloned into the primary vector pBin19 and transferred to Agrobacterium tumifaciens (LBA4404) by transformation and from there to potato tuber discs by

the method of Spychalla and Bevan (1993) Plant Tissue Culture Manual BII.

Example 3 - Transformation and overexpression constructs

In other transformants, full length cDNA clones encoding the isoamylase type of debranching enzyme in potato are used to increase debranching enzyme activity levels in transgenic potatoes. This is achieved by cloning each of them between the 2 x CaMV 35S promoter and the CaMV Terminator of pJIT60. Thence into a binary vector such as pBin19 between the T-DNA borders in *E. coli*. It is then transferred to Agrobacterium tumifaciens for transformation into plants.

Example 4 - Enzyme purification

Each different cDNA may be expressed in *E. coli* to define its activity closely and to obtain enough purified protein to produce an antiserum. This could be done using any suitable system e.g. the pSTAG expression vector for E. coli (strain K38)(Moyano et al (1996) Plant Cell 8: 1519-1532):

All three cDNA's were inserted into a vector which permitted expression in E.coli of proteins fused to a 15 amino acid tag at the N-terminus (pET Expression system, Novagen). The amount of expressed protein in E.coli extracts was then quantified by an assay for the S-tag and the proteins were then purified on an affinity matrix specific for the S-tag, The C21 and C15 cDNA's were inserted into the vector after removal of the fragment encoding the putative transit peptide (pET system manual 7th Edition, Novagen).

Fusion proteins from the C15 and C21 cDNA's were successfully expressed to high levels in E.Coli as

determined by using SDS-polyacrylamide gels (not shown). In both cases, single bands of protein not present in E.coli transformed with the vector alone are seen in crude extracts of both the soluble and unsoluble (inclusion body) fractions of the bacteria. These bands correspond closely in size to the predicted size for the expressed proteins: 87kD for C15 and 97kD for C21.

The recombinant isoamylase is most readily assayed when it is either purified from other hydrolases or by use of specific inhibitors to negate the contribution of interfering enzymes. It may also be visualised on non-denaturing glucan-containing polyacrylamide gels on which activities of starch hydolysing enzymes are separated and then revealed by staining of hydrolysis products with iodine (see Kakefuda et al, 1996 Planta 168: 175-182).

Example 5 - Preparation of antisera

Polyclonal antibodies against C15 and C21 were produced in New Zealand white rabbits using standard immunisation procedures (Harlow E. & Land D. (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbour, N.Y).

The immunoblot analysis was performed according to standard procedures (Sambrook J et al (1989) Molecular Cloning: A Laboratory Manual, Cold Spring harbor Laboratory Press, Cold Spring Harbor, NY). Filters were incubated with the rabbit antiserum and immunoreactive bands were detected using the methods of Towbin H. Et al (1997) Proc. Natl. Acad. Sci. USA 76: 4350-4354.

The presence of high-titer antibodies in antisera that recognised the proteins was demonstrated by immunoblot analysis of extracts from E.coli expressing either C15 or C21. Both C15 and C21 antisera immunoreact against C15

and C21.

Literature cited

Ball et al (1996) Cell 86, 349-352.

James et al (1995) Plant Cell 7, 417-429.

Mouille et al (1996) Plant Cell 8, 1353-1356.

Nakamura et al (1996a) Physiol. Plant 97, 491-498.

Nakamura et al (1996b) Planta 199, 209-218.

Pan and Nelson (1984) Plant Physiol. 74, 324-328.

CLAIMS

1. An isolated nucleic acid which comprises a nucleotide sequence which:
(a) encodes a polypeptide which has the properties of an isoamylase, and
(b) is obtainable from Solanum tuberosum

2. A nucleic acid as claimed in claim 1 wherein the polypeptide has an amino acid sequence shown in any one of Seq ID Nos 4, 5 or 6.

3. A nucleic acid as claimed in claim 1 or 2 comprising a nucleotide sequence encoding an isoamylase, the sequence:
(a) consisting of any one of Seq ID No 1, 2 or 3, or
(b) being degeneratively equivalent to any one of Seq ID Nos 1, 2 or 3.

4. A nucleic acid comprising a nucleotide sequence which:
(a) is a homologous variant of Seq ID No 1, 2 or 3, sharing at least 85, 90, 95, 96, 97, 98, or 99% sequence identity with any of said sequences, and
(b) encodes a polypeptide having isoamylase activity.

5. A nucleic acid as claimed in claim 5 wherein the variant is an allelic variant of Seq ID No 1, 2 or 3.

6. A nucleic acid as claimed in claim 4 wherein the variant is a derivative of Seq ID No 1, 2 or 3 by way of one or more of addition, insertion, deletion or substitution of one or more nucleotides.

7. A nucleic acid as claimed in claim 6 wherein the variant does not encode a transit peptide sequence.

8. A nucleic acid which is complementary to the nucleic acid of any one of claims 1 to 7.

9. A method for identifying or cloning an isoamylase from a plant species, which method employs a nucleic acid molecule having a nucleotide sequence of at least 15 nucleotides, which sequence is shown in, or complementary to, all or part of Seq ID No 1, 2 or 3.

10. A method as claimed in claim 9 comprising the steps of: (a) providing a preparation of nucleic acid from a plant cell,
(b) providing a nucleic acid molecule having a nucleotide sequence of at least 15 nucleotides, which sequence is shown in, or complementary to, all or part of Seq ID No 1, 2 or 3,
(c) contacting nucleic acid in said preparation with said nucleic acid molecule under conditions for hybridisation of said nucleic acid molecule to any nucleic acid encoding an isoamylase in said preparation,
(d) identifying said nucleic acid encoding an isoamylase if present by its hybridisation with said nucleic acid molecule, and optionally
(e) confirming the identity the isoamylase encoded by the nucleic acid by expressing it and assessing its activity.

11. A method as claimed in claim 10 wherein the hybridisation conditions are selected such to allow the identification of sequences having about 85% or more sequence identity with the nucleic acid molecule.

12. A method as claimed in claim 11 comprising use of two primers to amplify a nucleic acid encoding an isoamylase, at least one of the primers having a nucleotide sequence of at least 15 nucleotides, which sequence is shown in, or complementary to, all or part of Seq ID No 1, 2 or 3.

13. A method as claimed in claim 12 comprising the steps of:
(a) providing a preparation of nucleic acid from a plant cell,
(b) providing a pair of nucleic acid molecule primers suitable for PCR, at least one of the primers having a nucleotide sequence of at least 15 nucleotides, which sequence is shown in, or complementary to, all or part of Seq ID No 1, 2 or 3,
(c) contacting nucleic acid in said preparation with said primers under conditions for performance of PCR,
(d) performing PCR and determining the presence or absence of an amplified PCR product, and optionally
(e) confirming the identity of the amplified PCR product by expressing it and assesssing its isoamylase activity.

14. A nucleic acid molecule for use as a probe or primer in the method of any one of claims 10 to 13, said molecule having a nucleotide sequence of at least 15, 18, 21, 24 or 30 nucleotides, which sequence is shown in, or complementary to, all or part of Seq ID No 1, 2 or 3.

15. A recombinant vector comprising the nucleic acid of any one of claims 1 to 8.

16. A vector as claimed in claim 15 which is capable of replicating in a suitable host.

17. A vector as claimed in claim 15 or claim 16 wherein the nucleic acid is operably linked to a promoter or other regulatory element for transcription in a host cell.

18. A vector as claimed in claim 17 further comprising any one or more of the following: a terminator sequence; a polyadenylation sequence; an enhancer sequence; a marker gene.

19. A vector as claimed in claim 17 or claim 18 wherein the promoter is an inducible promoter.

20. A vector as claimed in any one of claims 15 to 19 which is a plant vector.

21. A vector as claimed in claim 20 comprising a selectable genetic marker which confers a selectable phenotype selected from: resistance to antibiotics or herbicides.

22. A method comprising the step of introducing a vector as claimed in any one of claims 15 to 21 into a cell.

23. A method for transforming a plant cell, comprising a method as claimed in claim 22, and further comprising the step of causing or allowing recombination between the vector and the plant cell genome to introduce the nucleic acid into the genome.

24. A host cell comprising a vector as claimed in any one of claims 15 to 21.

25. A host cell transformed with a vector as claimed in any one of claims 15 to 21.

26. A host cell as claimed in claim 24 or claim 25 which is a plant cell.

27. A host cell as claimed in claim 26 which is derived from a starch-synthesising organ.

28. A host cell as claimed in claim 27 which is in a tuber.

29. A host cell as claimed in any one of claims 26 to 28 which is in a plant.

30. A method for producing a transgenic plant comprising a method as claimed in claim 23 and further comprising the step of regenerating a plant from the transformed cell.

31. A plant comprising the cell of any one of claims 26 to 29.

32. A plant as claimed in claim 31 produced by the method of claim 30.

33. A plant which is the progeny of a plant as claimed in claim 31 or claim 32.

34. A plant as claimed in any one of claims 31 to 33 which is selected from: potato; pea; modified maize, wheat, cassava, rice and barley.

35. A part or propagule of the plant of any one of claims 31 to 34.

36. A polypeptide encoded by the nucleic acid of any one of claims 1 to 7.

37. A method of producing a polypeptide comprising the step of causing or allowing the expression from a nucleic acid of any one of claims 1 to 7 in a suitable host cell.

38. A composition comprising the polypeptide of claim 36.

39. An antibody or fragment thereof, or a polypeptide comprising the antigen-binding domain of the antibody, capable of specifically binding the polypeptide of claim 36.

40. A method of producing the antibody or fragment as claimed in claim 39 comprising the step of immunising a mammal with a polypeptide of claim 36.

41. A method of identifying and/or isolating an isoamylase comprising the step of screening candidate polypeptides with a polypeptide comprising the antigen-binding domain of the antibody of claim 39.

42. A method for the synthesis of a branched polysaccharide comprising the use of the polypeptide of claim 36.

43. A method for altering the quality or quantity of a polysaccharide in a host cell by influencing the isoamylase activity in that cell, the method comprising use of any one or more of the following: all or part of the nucleic acid of any one of claims 1 to 8; the polypeptide of claim 36; the antibody or fragment or polypeptide comprising the antigen-binding site thereof of claim 39.

44. A method as claimed in claim 43 wherein the subcellular location of the isoamylase activity is manipulated.

45. A method as claimed in claim 43 wherein the activity of two or more isoamylases is manipulated.

46. A method as claimed in any one of claims 42 to 45 wherein the polysaccharide is amylopectin.

47. A method as claimed in any one of claims 43 to 46 wherein the quality altered is the branching of the amylopectin.

48. A method as claimed in claim 47 wherein the

amylopectin is altered in at least one of the following ways:

(a) the degree of branching is decreased, or

(b) the degree of branching is increased, or

(c) the branching pattern is changed.

49. A method as claimed in any one of claims 43 to 48 comprising the step of causing or allowing expression of a nucleic acid according to any one of claims 1 to 8 within the cell.

50. A method as claimed in any one of claims 43 to 49 comprising repressing the isoamylase activity in the cell.

51. A method as claimed in claim 50 comprising the step of causing or allowing the transcription of part of the nucleic acid of any one of claims 1 to 7 in the cell such as to co-suppress the expression of an isoamylase.

52. A method as claimed in claim 50 comprising the step of causing or allowing the transcription of nucleic acid of claim 8 in the cell.

53. A method as claimed in claim 50 comprising the step of causing or allowing the expression of a polypeptide comprising the antigen-binding domain of the antibody of claim 39.

54. A method as claimed in any one of claims 43 to 53 wherein the cell is the plant cell of any one of claims 26 to 29.

55. A starch the quality of which has been altered in accordance with the method of any one of claims 43 to 54.

56. A plant product derived from any one of the plants

of claims 31 to 34 or the plant cells of claims 26 to 29, said product comprising a starch of claim 55.

57. A commodity comprising the starch of claim 55.

58. A commodity as claimed in claim 57 which is selected from: a human or animal foodstuff; a biodegradable plastic; a food-processing thickener; a starch coated film; a starch coated paper; a starch coated textile; a paint thickener; a mining explosive; a pharmaceutical; a glue.

Fig. 1A


```
   1 CTTATGGGAC TTGATTAAGA ATATGTGATC CACCAAGTTC TATATCTGAC
  51 GCTGTCGTAA CATTGTGTGC TGCTAATGGC AACTTCACCA ATACAGTTGG
 101 CTGTGCATTC ACGTTTGTTG AGCTATGGCA GTACTGAGTC AACCAAGTTG
 151 GTTCCTTCAT CATCAGGTAA CCGTGGAAAA ATAGTATGCA GTCTAAGGAA
 201 GCTGGAATTG GAAGACATGA ATTTCTCTGG CATAGGTCGA AATAATGATC
 251 AAGAAGCTCC TAGGAGAGCT CATCGACGAA AAGCACTATC AGCATCGAGA
 301 ATTTCGCTTG TTCCATCTGC AAAAAGGGTT CCCACTTACC TTTTCAGGAC
 351 AGATATTGGT GGTCAAGTGA AAGTCTTGGT GGAAAGGACA AATGGAAAGT
 401 ACAAAGTGCT TGTAGAAGTA TTGCCATTGG AGCTCTCATA TGCACATTCT
 451 GAGCTGGTTA TGGTTTGGGG TCTTTTTAGA TCTGATGCTT CATGCTTTAT
 501 GCCTCTAGAT CTAAATAGAC GTGGAGCAGA TGGAAAAAGT AGTACTGTTG
 551 AAACACCATT TGTGCAAGGA CCTTCAGGCA AGGTCACCGT GGAGCTGGAT
 601 TTTGAAGCAA GTTTAGCCCC CTTCTATATC TCCTTCTATA TGAAGTCGCA
 651 ACTAGTTTCT GACATGGAAA ACTCAGAAAT CAGAAGTCAC AGGAACACAA
 701 ATTTTGTTGT ACCAGTTGGT CTCAGTTCAG GGCATCCTGC TCCATTGGGT
 751 ATTTCCTTTC AGCCAGATGG ATCTGTGAAT TTTGCTCTCT TCTCACGCAG
 801 TGCAAGAAGT GTAGTTCTGT GCTTGTATGA TGACATATCA GTTGAAAAAC
 851 CTTCTTTAGA GATTGATCTA GATCCTTATA TTAATCGATC AGGCGATATT
 901 TGGCATGCTG CTTTAGATTG TTCTTTGCCA TTTAAGACTT ATGGTTATAG
 951 ATGTAAGGCG ACTACTTCTG GGAAGGGAGA GCTGGTTCTT TTGGACCCAT
1001 ATGCTAAGGT GATAAGGCGT GTTATTCCTC GTCAGGGTGG GTCTGAGATA
1051 CGTCCAAAAT ATCTTGGAGA ACTATGCCTG GAACCTGGCT ATGATTGGAG
1101 CGGTGATGTC CCCCCTAGCT TACCTATGGA GAAACTAATA ATTTACCGCT
1151 TAAATGTGAC TCAATTTACA AAGGACAAGT CCAGTAAGCT ACCTGATGAC
1201 CTTGCTGGAA CTTTCTCTGG CATTAGCGAA AAATGGCACC ATTTTAAAGA
1251 TCTTGGTGTG AATGCAATGT TACTGGAGCC AATTTTCCCT TTTGATGAGC
1301 AGAAAGGACC CTATTTTCCG TGGCATTTCT TCTCACCTGG AAATATGTAT
1351 GGACCTTCTG GTGACCCTCT TTCTGCCATT AAATCGATGA AGGATATGGT
1401 TAAGAAATTA CATGCTAACG GGATAGAGGT TTTTCTTGAA GTTGTTTTCA
1451 CTCACACTGC AGAGGATGCA CCTTTGATGA ATGTTGATAA CTTTTCATAT
1501 TGCATAAAAG GTGGTCAGTA TCTGAATATT CAAAATGCAT TGAATTGCAA
```

1551 TTACCCCATA GTCCAACAAA TGATTTTGGA CTGTCTCCGC CACTGGGTAA

1601 TTGAGTTTCA TATTGATGGT TTTGTTTTTG TCAACGCTTC TTCCTTGTTG

1651 AGAGGGTTCA ATGGAGAGAT TCTATCTCGT CCTCCATTAG TTGAAGCTAT

1701 TGCCTTTGAT CCTATCCTTT CAAAGGTCAA GATGATTGCA GATAATTGGA

1751 ATCCATTAAC CAATGATTCG AAGGAAAATT TATTCCCTCA CTGGAGGAGA

1801 TGGGCAGAGA TAAATATGAG ATTTTGTGAT GACATTCGAG ACTTCTTGAG

1851 AGGCGAGGGT CTTCTAAGCA ATCTAGCAAC ACGACTTTGT GGAAGTGGGG

1901 ATATCTTCGC AGGTGGACGT GGTCCTGCAT TCTCTTTTAA TTATATTGCC

1951 AGAAATTCTG GACTCACACT TGTTGACCTA GTTAGCTTCA GTAGTAATGA

2001 AGTGGCTTCA GAGTTAAGTT GGAACTGTGG ACAAGAAGGC GCTACGACCA

2051 ATAACATTGT CCTAGAGAGA CGACTTAAAC AAGTTCGTAA TTTTCTGTTC

2101 ATATTGTTCA TTTCTCTAGG TGTACCAGTA CTTAACATGG GAGACGAGTG

2151 TGGTCAGTCT TCAGGAGGTC CCCCTGCATA TGATGCTCGA AAATCTTTGG

2201 GTTGGAATAC TTTAAAAACT GGTTTTGGGA CTCAGATTGC CCAGTTTATT

2251 TCATTCTTGA GTAATTTAAG AATGAGAAGA AGTGATCTTC TTCAAAAGAG

2301 AACCTTCTTG AAGGAAGAAA ACATCCAGTG GCATGGGAGT GACCAATCTC

2351 CTCCGAAATG GGATGGCCCG TCTAGCAAAT TCTTGGCTAT GACTTTGAAG

2401 GCCGATGCTG AAGTCAGCCA GACATTAGTC TCTGATATCG TAGGTGACCT

2451 GTTTGTTGCT TTCAATGGTG CTGGTGATTC AGAGATTGTT ATCCTTCCAC

2501 CTCCTCCAAC AGATATGGTA TGGCATCGTC TCGTTGACAC AGCCCTCCCT

2551 TTCCCGGGGT TTTTCGATGA GAAGGGAACT CCAGTTGAAG ATGAATTAGT

2601 TGCTTATGAG ATGAAGTCTC ACAGCTGTTT GCTGTTTGAA GCTCAGAGAC

2651 TAGCTGAAAT AGATTCTAGC AAGAGAAAGA AACAGATTAG ACTTTCTTCT

2701 AAGAGGCAAT AGTTTGTAAA GCCCCTAAGT ATATATATAT GTTTAAATAA

2751 GAGGCTTTTT TTTCTGAATA AATAAGAAGA TTTTACTGAG AATACTTGTA

2801 TCTAAACATT TTCTTTTGCA GCTTCAAATA AAAAAAAAAA AAA

Fig. 1B

Fig. 2A

```
   1 CTCAGTCCTT CTCAATTTCA GTGCCACATA CTCTAGATCA CACTCTCTCT
  51 TCTTCCTCAA AGTTCTCCCA TGGAGTTACT TCATTGTCCT TCCATTTCTA
 101 CCTACAAACC TAAACTCTCT TTCCACAACC ATCTTTTCTC GAGGAGAAGC
 151 AGTAACGGTG TAGATTTTGA GAGTATTTGG AGAAAATCGA GGTCTTCAGT
 201 GGTTAATGCT GCTGTTGATA GTGGACGTGG AGGTGTGGTG AAGACTGCGG
 251 CTACTGCGGT GGTGGTGGAG AAGCCGACGA CGGAACGATG TCGTTTTGAG
 301 GTTTTATCAG GGAAGCCATT GCCGTTTGGT GCTACTGCGA CAGATGGTGG
 351 TGTGAATTTC GCTGTTTTTT CAAGGAATGC TACAGCTGCT ACTCTTTGCT
 401 TGATCACTCT TTCCGATTTA CCTGAGAAGA GAGTGACCGA GCAAATTTTC
 451 CTGGATCCTC TAGCTAATAA AACTGGAGAT GTATGGCATG TGTTCCTTAA
 501 GGGAGATTTT GAGAATATGC TATATGGCTA CAAATTTGAT GGGAAATTCT
 551 GTCCTGAAGA AGGACACTAC TTTGACTCTT CGCAGATAGT GTTGGATCCT
 601 TATGCCAAGG CTATAGTAAG CAGAGGAGAA TATGGTGTAT TAGGGCCAGA
 651 GGATGATTGT TGGCCCCCAA TGGCTGGCAT GGTACCTTCT GCTTCTGATC
 701 AGTTTGATTG GGAAGGAGAT CTACCACTGA AGTTTCCACA GAGAGATCTT
 751 GTAATCTATG AAATGCATGT TCGTGGGTTT ACTAATCATG AGTCGAGTGA
 801 AACAAAATAT CCTGGTACTT ACCTTGGTGT TGTGGAGAAA CTTGATCACT
 851 TGAAGGAACT TGGTGTCAAC TGTATAGAGC TAATGCCCTG TCACGAGTTC
 901 AATGAGCTGG AGTACTATAG TTATAACTCT GTATTGGGCG ACTACAAGTT
 951 TAACTTTTGG GGCTATTCTA CTGTCAATTT CTTTTCTCCA ATGGGAAGAT
1001 ACTCATCTGC TGGTCTAAGT AATTGCGGCC TCGGTGCAAT AAACGAATTT
1051 AAGTATCTTG TCAAGGAAGC ACATAAACGT GGAATCGAGG TTATCATGGA
1101 TGTTGTTTTC AATCACACTG CTGAAGGAAA TGAAAATGGT CCCATACTAT
1151 CATTTAGAGG CATTGACAAC AGTGTGTTTT ATACGCTAGC TCCTAAGGGT
1201 GAATTTTACA ACTACTCAGG ATGTGGAAAT ACCTTCAACT GTAATAATCC
1251 CATTGTACGT CAATTTATAG TGGATTGCTT GAGATATTGG GTTACCGAAA
1301 TGCACGTAgA TGGCTTCCGC TTTGATCTTG CTTCTATCCT TACAAGAAGT
1351 AGCAGCTCGT GGAATGCTGT AAATGTCTAT GGAAATTCAA TTGACGGTGA
1401 CGTGATCaCC ACAGGCACTC CTCTCACAAG CCCACCATTG ATTGATATGA
1451 TTAGCAATGA TCCAATACTT CGTGGAGTAA AGCTTATAGC TGAAGCATGG
1501 GATTGTGGAG GCCTTTACCA AGTTGGCATG TTTCCGCACT GGGGTATCTG
```

```
1551  GTCGGAGTGG AACGGAAAGT ACCGTGACAT GGTACGGCAG TTCATCAAAG
1601  GCACTGATGG GTTTTCTGGG GCTTTTGCTG AATGCCTTTG TGGAAGCCCA
1651  AATCTATACC AGAAAGGAGG AAGAAAACCA TGGAACAGTA TAAATTTCGT
1701  GTGTGCCCAC GATGGTTTTA CTTTGGCTGA TTTAGTGACA TACAACAATA
1751  AACACAATTT GGCAAATGGA GAGGACAACA AAGACGGGGA GAATCACAAT
1801  AATAGTTGGA ATTGTGGTGA GGAAGGAGAA TTTGCAAGTA TCTTTGTGAA
1851  GAAATTGAGG AAAAGACAAA TGCGGAACTT CTTCcTCTGC cTTAtGGTTT
1901  CCCAAGGTGT TCCCATGATA TATATGGGCG ATGAATATGG TCACACTAAG
1951  GGAGGAAACa ACAACACGtA TTGCCATGAT AATTATATTA ATTACTTCCG
2001  TTGGGATAAG AAGGATGAAT CTTCATCTGA TTTTTTGAGA TTTTGCGGCC
2051  TCATGACCAA ATTCCGCCAT GAATGTGAAT CACTGGGATT AGATGGTTTC
2101  CCTACAGCAG AAAGGCTGCA ATGGCATGGT CACACTCCTA GAACTCCAGA
2151  TTGGTCTGAA ACAAGTCGAT TCGTTGCATT CACACTGGTC GACAAAGTGA
2201  AGGGAGAACT ATATATTGCC TTTAACGCCA GCCATTTGCC TGTAACGATT
2251  ACACTTCCAg ATAGGCCTGG TTATAGATGG CAGCCGTTTG TGGACACAGG
2301  CAAACCAGCA CCATTTGACT TCTTGACAGA CGACGTTCCT GAGAGAGAGA
2351  CAGCAGCCAA ACAATATTCT CATTTTCTGG ACGCGAACCA GTATCCGATG
2401  CTCAGTTATT CATCCATTAT TCTTTTACTA TCATCTGCTG ATGATGCATA
2451  GTTTCATTCA CCAAGTTAGG TGGAGGTAAA TCAGCTTCAG ATTTTGTTAT
2501  ATGCAGTGAG GTGTTACTTT GTAAATAAAA GTAAGAAGCA GGACAGAACA
2551  GAACTGCAAA CGGATAAAAT TTGTGAGGAA GAAGCTGATG ATTTATAAGA
2601  tACACCTTGT aTTtTAATtG CATTTATATA AAATAAAATA nTAGTGAAAT
2651  TGTcTGTGcG AAAaaaaaaa AAAAAAAAAA TAAAAAAAAA AAAAAAAAAA
2701  AAAAAA
```

Fig. 2B

Fig. 3A

```
   1 CACAGATTCT CTTCTCCAAA AATAGGGCCC GATGATTAGA GGACCACCAC
  51 AAATCGTCCA GAAATGTCCA ACCGACATTG TAACAGTTAA CCGGACCAAT
 101 ATAGTTCCAC GAACGCACCG TCACGCTCTT CAAGATCTCC GGCAGCTTCG
 151 CCGGCGTGAT AGTCTCAGGC TCTTCTCCTC TGATCACCGG ATTCTGAAGT
 201 TTTGTACATC GGAGGAGGCG TTCCAACCTA GGTTGGTCGC AGCAGCTAAA
 251 CTTCAGGAAG AAGCTCCTCA AATGCTGGAC ACTTTCCCTT CATTCAAAGT
 301 TTCCCCTGGT CTGGCTCATC CACTAGGAGT ATCAGAAACT GAAAGTGGAA
 351 TAAATTTTGC AATTTTTTCT CAGCATGCTT CTGCAGTTAC ACTTTGCATA
 401 ATTCTTCCAA AGAGTGTTCA TGATGGAATG ATTGAATTAG CATTGGATCC
 451 ACAGAAGAAC CGCACAGGAG ACATATGGCA CATATGCATT AAGGAGTTGC
 501 CCCAAGGTGG TGTCCTTTAT GGTTATCGCA TTGATGGACC TCGAAATTGG
 551 CATGAAGGGC ATCGATTTGA TGATAGCATT ATTTTGGTTG ATCCTTACGC
 601 AAAACTAATT GAAGGTCGAC GAGTTTTTGG AGATGAAAGC AATAAAATGT
 651 GTAGATTTTT TGGAACTTAT GATTTCAATA GCTTGCCTTT TGACTGGGGA
 701 GAAAATTACA AGCTTCCAAA TATACCCGAG AAAGATCTTG TTATATATGA
 751 GATGAATGTT CGTGCTTTTA CTGCTGATGA AACAAGTAGT TTGGATCAAG
 801 ATCAACGGGG AAGTTACCTT GGCTTAATTG AAAAGATACC ACATCTTCTC
 851 GAGCTTGGTG TCAATGCAGT AGAATTATTG CCTGTTTTTG AGTTTGATGA
 901 ACTGGAATTA CAAAGGCGAC CTAATCCGAG AGATCACATG ATCAATACAT
 951 GGGGCTACTC AACAATAAAC TTTTTTGCTC CAATGAGTCG ATATGCAAGT
1001 TGTGGTGGCG GACCTGTCCG TGCTTCCTGG GAGTTCAAAG AAATGGTCAA
1051 GGCCTTGCAT GGTGCTGGAA TTGAGGTCAT CTTAGATGTT GTTTATAATC
1101 ACACAAATGA AGCTGATGAT GAAAACCCAT ATACAACCTC ATTCCGAGGA
1151 ATAGACAACA AGGTTTATTA CATGGTAGAT TTAAACAACA ATGCTCAGCT
1201 GCTGAATTTC GCTGGATGTG GAAATACTTT TAACTGCAAC CATCCcACAG
1251 TCATGGAACT TATACTTGAA AGCTTAAGAC ACTGGGTCAC CGAGTATCAT
1301 GTCGATGGAT TTCGCTTTGA TCTTGCTAGT GTTCTTTGCA GAGGGACAGA
1351 TGGTACTCCC ATTAATGCTC CCCCCCTTGT TAAGGCCATT TCCAAAGATA
1401 GTGTATTGTC GAGGTGCAAA ATTATTGCTG AGCCATGGGA TTGTGGAGGC
1451 CTATATCTTG TTGGAAAGTT TCCGAACTGG GACCGGTGGG CTGAGTGGAA
1501 TGGGAAGTAC CGCGATGACA TCAGGAGATT TATAAAGGGC GATGCTGGCA
```

1551 TGAAAGGAAA TTTTGCAACC CGTATCGCAG GTTCAGCGGA TCTGTACAGA

1601 GTGAACAAGC GAAAGCCGTA CCACAGTGTC AACTTCGTGA TTGCCCATGA

1651 TGGCTTTACC TTGTATGACC TTGTTTCATA CAATAATAAG CACAATGATG

1701 CAAACGGTGA AGGTGGCAAT GATGGATGCA ATGACAACTT CAGTTGGAAT

1751 TGTGGAATTG AAGGTGAAAC TTCAGATGCA AATATTAACG CACTGCGTTC

1801 ACGGCAAATG AAAAATTTTC ATTTGGCACT GATGGTTTCT CAGGGAACAC

1851 CAATGATGCT TATGGGGGAT GAGTATGGGC ATACCCGCTA TGGAAATAAT

1901 AACAGTTATG GACATGATAC CGCCATCAAC AATTTCCAGT GGGGACAATT

1951 GGAAGCAAGG AAGAATGATC ACTTCAGGTT CTTTTCCAAG ATGATAAAGT

2001 TTCGACTGTC CCACAATGTT CTTAGAAAGG AAAACTTCAT TGAGAAGAAC

2051 GACATTACCT GGCTCGAGGA CAACTGGTAC AATGAAGAGA GTAGATTCCT

2101 TGCATTTATG CTCCATGATG GGAATGGAGG AGATATTTAC TTGGCATTTA

2151 ATGCACACCA CTTCTCCATC AAAACAGCAA TACCTTCACC ACCACGAAAT

2201 AGAAGTTGGT ACCGAGTGGT GGACACTAAT CTGAAATCAC CAGATGATTT

2251 TGTTACTGAG GGAGTGTCTG GTATCAGTAA AACTTATGAT GTTGCGCCGT

2301 ACTCTGCTAT CCTTCTTGAA GCAAAGCAAT AATTACCGGG ACTATGCTGC

2351 TTTAGATGTT GTCCATGAGT TATTACAGTA TTACCTCCTT CTGGATTGGA

2401 TAGTTCAAAT CGGAATTCAG GCTGTTAGCC TATAGATGTT TGCAATAAGC

2451 AACCAGTTTG TTCAAGCTGC TATTGACAGG TACAAACACC CCATAGTAAT

2501 AAGATAAACT GAGACCATTG ATCCAAAAAA AAAAAAAAAA AAAAAAAAAA

2551 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

2601 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAA

Fig.3B

```
  1  MATSPIQLAV HSRLLSYGST ESTKLVPSSS GNRGKIVCSL RKLELEDMNF
 51  SGIGRNNDQE APRRAHRRKA LSASRISLVP SAKRVPTYLF RTDIGGQVKV
101  LVERTNGKYK VLVEVLPLEL SYAHSELVMV WGLFRSDASC FMPLDLNRRG
151  ADGKSSTVET PFVQGPSGKV TVELDFEASL APFYISFYMK SQLVSDMENS
201  EIRSHRNTNF VVPVGLSSGH PAPLGISFQP DGSVNFALFS RSARSVVLCL
251  YDDISVEKPS LEIDLDPYIN RSGDIWHAAL DCSLPFKTYG YRCKATTSGK
301  GELVLLDPYA KVIRRVIPRQ GGSEIRPKYL GELCLEPGYD WSGDVPPSLP
351  MEKLIIYRLN VTQFTKDKSS KLPDDLAGTF SGISEKWHHF KDLGVNAMLL
401  EPIFPFDEQK GPYFPWHFFS PGNMYGPSGD PLSAIKSMKD MVKKLHANGI
451  EVFLEVVFTH TAEDAPLMNV DNFSYCIKGG QYLNIQNALN CNYPIVQQMI
501  LDCLRHWVIE FHIDGFVFVN ASSLLRGFNG EILSRPPLVE AIAFDPILSK
551  VKMIADNWNP LTNDSKENLF PHWRRWAEIN MRFCDDIRDF LRGEGLLSNL
601  ATRLCGSGDI FAGGRGPAFS FNYIARNSGL TLVDLVSFSS NEVASELSWN
651  CGQEGATTNN IVLERRLKQV RNFLFILFIS LGVPVLNMGD ECGQSSGGPP
701  AYDARKSLGW NTLKTGFGTQ IAQFISFLSN LRMRRSDLLQ KRTFLKEENI
751  QWHGSDQSPP KWDGPSSKFL AMTLKADAEV SQTLVSDIVG DLFVAFNGAG
801  DSEIVILPPP PTDMVWHRLV DTALPFPGFF DEKGTPVEDE LVAYEMKSHS
851  CLLFEAQRLA EIDSSKRKKQ IRLSSKRQ
```

Fig. 4

```
  1  MELLHCPSIS TYKPKLSFHN HLFSRRSSNG VDFESIWRKS RSSVVNAAVD
 51  SGRGGVVKTA ATAVVVEKPT TERCRFEVLS GKPLPFGATA TDGGVNFAVF
101  SRNATAATLC LITLSDLPEK RVTEQIFLDP LANKTGDVWH VFLKGDFENM
151  LYGYKFDGKF CPEEGHYFDS SQIVLDPYAK AIVSRGEYGV LGPEDDCWPP
201  MAGMVPSASD QFDWEGDLPL KFPQRDLVIY EMHVRGFTNH ESSETKYPGT
251  YLGVVEKLDH LKELGVNCIE LMPCHEFNEL EYYSYNSVLG DYKFNFWGYS
301  TVNFFSPMGR YSSAGLSNCG LGAINEFKYL VKEAHKRGIE VIMDVVFNHT
351  AEGNENGPIL SFRGIDNSVF YTLAPKGEFY NYSGCGNTFN CNNPIVRQFI
401  VDCLRYWVTE MHVDGFRFDL ASILTRSSSS WNAVNVYGNS IDGDVITTGT
451  PLTSPPLIDM ISNDPILRGV KLIAEAWDCG GLYQVGMFPH WGIWSEWNGK
501  YRDMVRQFIK GTDGFSGAFA ECLCGSPNLY QKGGRKPWNS INFVCAHDGF
551  TLADLVTYNN KHNLANGEDN KDGENHNNSW NCGEEGEFAS IFVKKLRKRQ
601  MRNFFLCLMV SQGVPMIYMG DEYGHTKGGN NNTYCHDNYI NYFRWDKKDE
651  SSSDFLRFCG LMTKFRHECE SLGLDGFPTA ERLQWHGHTP RTPDWSETSR
701  FVAFTLVDKV KGELYIAFNA SHLPVTITLP DRPGYRWQPF VDTGKPAPFD
751  FLTDDVPERE TAAKQYSHFL DANQYPMLSY SSIILLLSSA DDA
```

Fig. 5

```
  1  MIRGPPQIVQ KCPTDIVTVN RTNIVPRTHR HALQDLRQLR RRDSLRLFSS
 51  DHRILKFCTS EEAFQPRLVA AAKLQEEAPQ MLDTFPSFKV SPGLAHPLGV
101  SETESGINFA IFSQHASAVT LCIILPKSVH DGMIELALDP QKNRTGDIWH
151  ICIKELPQGG VLYGYRIDGP RNWHEGHRFD DSIILVDPYA KLIEGRRVFG
201  DESNKMCRFF GTYDFNSLPF DWGENYKLPN IPEKDLVIYE MNVRAFTADE
251  TSSLDQDQRG SYLGLIEKIP HLLELGVNAV ELLPVFEFDE LELQRRPNPR
301  DHMINTWGYS TINFFAPMSR YASCGGGPVR ASWEFKEMVK ALHGAGIEVI
351  LDVVYNHTNE ADDENPYTTS FRGIDNKVYY MVDLNNNAQL LNFAGCGNTF
401  NCNHPTVMEL ILESLRHWVT EYHVDGFRFD LASVLCRGTD GTPINAPPLV
451  KAISKDSVLS RCKIIAEPWD CGGLYLVGKF PNWDRWAEWN GKYRDDIRRF
501  IKGDAGMKGN FATRIAGSAD LYRVNKRKPY HSVNFVIAHD GFTLYDLVSY
551  NNKHNDANGE GGNDGCNDNF SWNCGIEGET SDANINALRS RQMKNFHLAL
601  MVSQGTPMML MGDEYGHTRY GNNNSYGHDT AINNFQWGQL EARKNDHFRF
651  FSKMIKFRLS HNVLRKENFI EKNDITWLED NWYNEESRFL AFMLHDGNGG
701  DIYLAFNAHH FSIKTAIPSP PRNRSWYRVV DTNLKSPDDF VTEGVSGISK
751  TYDVAPYSAI LLEAKQ
```

Fig. 6

```
  1  AGTAGTTTTT ACACATACTG CTGATTCTGG AGCTCTTCGT GGAATTGATG
 51  ACAGTTCCTA TTACTACAAG GGAAGAGCCA ATNATCTAGA TTCTAAAAGT
101  TACTTGAACT GTAACTATCC TGTTGTTCAG CAGTTGGTAT TGGAGAGCTT
151  GCGTTATTGG GTAACCGAGT TTCATGTAGA TGGATTTTNT TTTATAAATN
201  CTTCATCTCT CTTGAGAGGC GTTCACGGTG AACAGCTCTC TCGTCCTCCT
251  TTGGTTGAAG CAATAGCTTT TNATCCACTT CTTGCGGAGA CCAAACTAAT
301  AGCTGATTGC TGGGNTCCAC TTGAAATGNT GCCANAAGAA GTACGGGTTC
351  CCACAATTTG GAAGCNATNG GCAGAACTCA NNNCAAGGTN TTTTTCGAAA
401  TNTNAGGAAA TTTTTTAAGG GGAANGGG
```

Fig. 7

```
1  MATSPIQLAV HSRLLSYGST ESTKLVPSSS GNRGKIVC
```

Fig. 8A

```
1  MELLHCPSIS TYKPKLSFHN HLFSRRSSNG VDFESIWRKS RSSVVNA
```

Fig. 8B

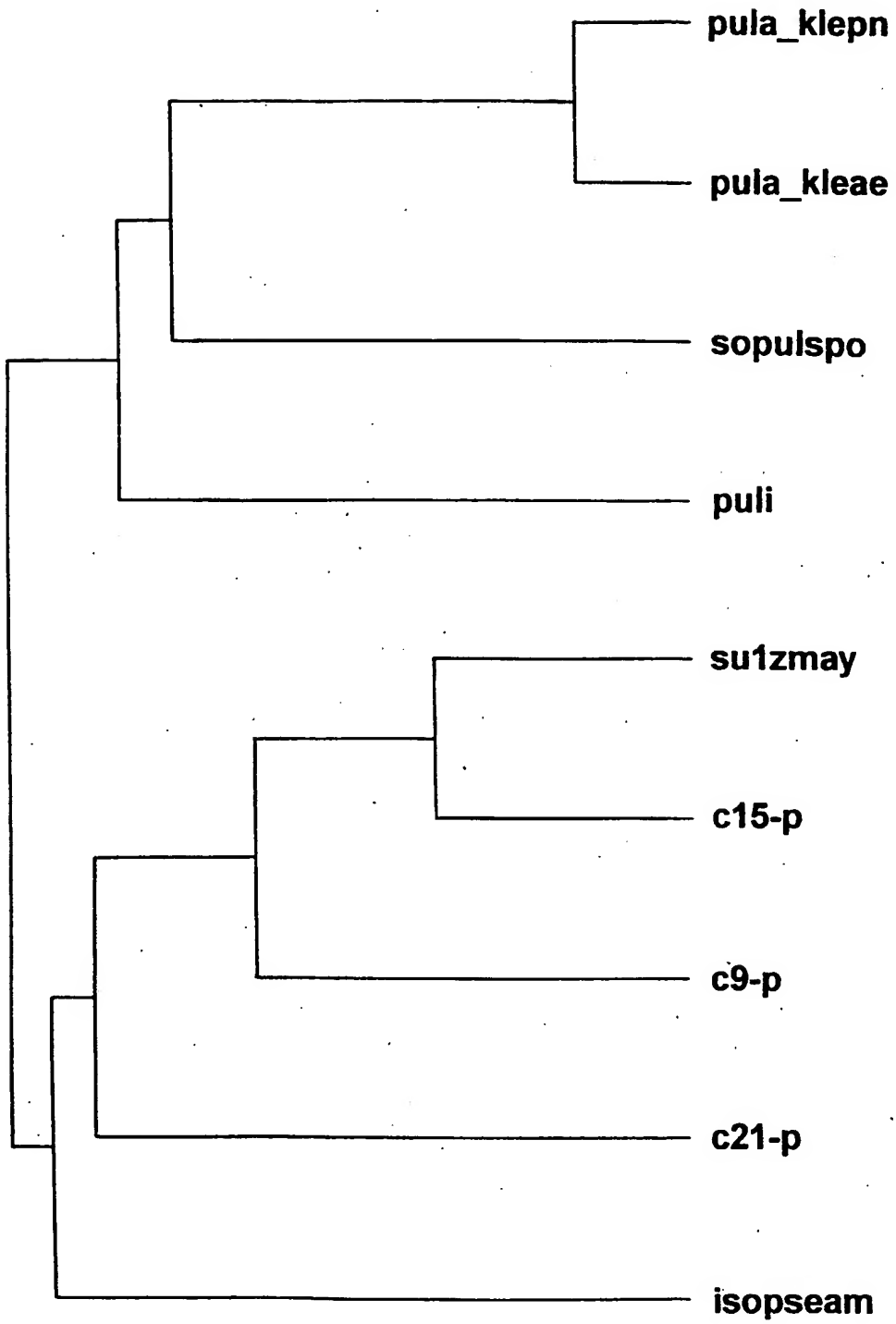
pula_klepn
pula_kleae
sopulspo
puli
su1zmay
c15-p
c9-p
c21-p
isopseam

Fig. 9

Fig. 10A

```
   1  CTTATGGGAC TTGATTAAGA ATATGTGATC CACCAAGTTC TATATCTGAC
  51  GCTGTTGTAA CATTGTGTGC TGCTAATGGC AACTTCACCA ATACAGTTGG
 101  CTGTGCATTC ACGTTTGTTG AGCTATGGCA GTACTGAGTC AACCAAGTTG
 151  GTTCCTTCAT CATCAGGTAA CCGTGGAAAA ATAGTATGCA GTCTAAGGAA
 201  GCTGGAATTG GAAGACATGA ATTTCTCTGG CATAGGTCGA AATAATGATC
 251  AAGAAGCTCC TAGGAGAGCT CATCGACGAA AAGCACTATC AGCATCGAGA
 301  ATTTCGCTTG TTCCATCTGC AAAAAGGGTT CCCACTTACC TTTTCAGGAC
 351  AGATATTGGT GGTCAAGTGA AAGTCTTGGT GGAAAAGACA AATGGAAAGT
 401  ACAAAGTGCT TGTAGAAgTC TTGCCATTGG AGCTCTCAGA TGCACATTCT
 451  GAGCTAGTTA TGGTTTGGGG TCTTTTTAGA TCTGATGCTT TATGCTTTAT
 501  GCCTCTGGAT CTAAACAGAC GTGGAGCAGA TGGAAAAAGT AGTACTGTTG
 551  AAACACCATT TGTGCAAGGA CCTTCAGGCA AGGTCACCGT GGAGCTGGAT
 601  TTTGAAGCAA GTTTAGCCCC CTTCTATATC TCCTTCTATA TGAAGTCACA
 651  ACTAGTTTCT GACATGGAAA ACTCAGAAAT CAGAAGTCAC AGGAACACAA
 701  ATTTTGTTGT ACCAGTTGGT CTCAGTTCAG GGCATCCTGC TCCATTGGGT
 751  ATTTCCTTTC AGCCAGATGG ATCTGTGAAT TTTGCTCTCT TCTCACGCAG
 801  TGCAAGAAGT GTAGTTCTGT GCTTGTATGA TGACATATCA GTTGAAAAAC
 851  CTTCTTTAGA GATTGATCTA GATCCTTATA TTAATCGATC AGGCGATATT
 901  TGGCATGCTG CTTTAGATTG TTCTTTGCCA TTTAAGACTT ATGGTTATAG
 951  ATGTAAGGCG ACTACTTCTG GGAAGGGAGA GCTGGTTCTT TTGGACCCAT
1001  ATGCTAAGGT GATAAGGCGT GTTATTCcTc GTCAGGGTGG GTCTGAGATA
1051  CGTCCAAAAT ATCTTGGAGA ACTATGCCTG GAACCTGGCT ATGATTGGAG
1101  CGGTGATGTC CCCCCTAGCT TACCTATGGA GAAACTAATA ATTTACCGCT
1151  TAAATGTGAC TCAATTTACA AAGGACAAGT CCAGTAAGCT ACCTGATGAC
1201  CTTGCTGGAA CTTTcTcTGG CATTAGCGAA AAATGGCACC ATTTTAAAGA
1251  TCTTGGTGTG AATGCAATGT TACTGGAGCC AATTTTCCcT TTTGATGAGC
1301  AGAAAGGACC CTATTTTCCG TGGCATTTCT TTTCACCTGG AAATATGTAT
1351  GGACCTTCTG GTGACCCTCT TTCTGCCATT AAATCGATGA AGGATATGGT
1401  TAAGAAATTA CATGCTAACG GGATAGAGGT TTTTCTTGAA GTTGTTTTCA
1451  CTCACACTGC AGAGGATGCA CCTTTGATGA ATGTTGATAA CTTTTCATAT
1501  TGCATAAAAG GTGGTCAGTA TCTGAATATT CAAAATGCAT TGAATTGCAA
```

1551 TTACCCCATA GTCCAACAAA TGATTTTGGA CTGTCTCCGC CACTGGGTAA

1601 TTGAGTTTCA TATTGATGGT TTTGTTTTTG TCAACGCTTC TTCCTTGTTG

1651 AGAGGGTTCA ATGGAGAGAT TCTATCTCGT CCTCCATTAg TTgaagcTaT

1701 TGCCTTTGAT CCTATCCTTT CAAAGGTCAA GATGATTGCA GATAATTGGA

1751 ATCCATTAAC CAATGATTCG AAGGAAAATT TATTCCCTCA CTGGAGGAGA

1801 TGGGCAGAGA TAAATATGAG ATTTTGTGAT GACATTCGAG ACTTCTTGAG

1851 AGGCGAgGGT CTTCTAAnCA ATCTAnCAAC ACgACTTTGT GGAAGTGGGG

1901 ATATCTTCGC AGGTGGACGT GGTCCTGCAT TCTCTTTTAA TTATATTGCC

1951 AGAAATTCTG GACTCACACT TGTTGACCTA GTTAGCTTCA GTAGTAATGA

2001 AGTGGCTTCA GAGTTAAGTT GGAACTGTGG ACAAGAAGGC GCTACGACCA

2051 ATAACATTGT CCTAGAGAGA CGACTTAAAC AAGTTCGTAA TTTTCTGTTC

2101 ATATTGTTCA TTTCTCTAGG TGTACCAGTA CTTAACATGG GAGACGAGTG

2151 TGGTCAGTCT TCAGGAGGTC CCCCTGCaTa TgATgCTCGA AAATCTTTGG

2201 GTTGGAATAC TTTAAAAACT GGTTTTGGGA CTCAGATTGC CCAGTTTATT

2251 TCATTCTTGA GTAATTTAAG AATGAGAAGA AGTGATCTTC TTCAAAAGAG

2301 AACCTTCTTG AAGGAAGAAA ACATCCAGTG GCATGGGAGT GACCAATCTC

2351 CTCCGAAATG GGATGGCCCG TCTAGCAAAT TCTTGGCTAT GACTTTGAAG

2401 GCCGATGCTG AAGTCAGCCA GACATTAGTC TCTGATATCG TAGGTGACCT

2451 GTTTGTTGCT TTCAATGGTG CTGGTGATTC AGAGATTGTT ATCCTTCCAC

2501 CTCCTCCAAC AGATATGGTA TGGCATCGTC TCGTTGACAC AGCCCTCCCT

2551 TTCCCGGGGT TTTTCGATGA GAAGGGAACT CCAGTTGAAG ATGAATTAGT

2601 TGCTTATGAG ATGAAGTCTC ACAGCTGTTT GCTGTTTGAA GCTCAGAGAC

2651 TAGCTGAAAT AGATTCTAGC AAGAGAAAGA AACAGATTAG ACTTTcTTcT

2701 AAGAGGCAAT AGTTTGTAAA GCCCCTAAGT ATATATATAT GTTTAAATAA

2751 GAGGCTTTTT TTTCTGAATA AATAAGAAGA TTTTACTGAG AATACTTGTA

2801 TCTAAACATT TTCTTTTGCA GCTTCAAATA AAAAAAAAAA AAA

Fig. 10B

Fig. 11A

```
   1  GGCGGCCGCT CTAGAACTAG TGGATCCCCC GGGCTGCAGG AATTCGAGGA
  51  TCCGGGTACC ATGGCTCAGT CCTTCTCAAT TTCAGTGCCA CATACTCTAG
 101  ATCACACTCT CTCTCTTCCT CAAAGTTCTC CCATGGAGTT ACTTCATTGT
 151  CCTTCCATTT CTACCTACAA ACCTAAACTC TCTTTCCACA ACCATCTTTT
 201  CTCGAGGAGA AGCAGTAACG GTGTAgATTT TGAGAGTATT TGGAGAAAAT
 251  CGAGGTCTTC AGTGGTTAAT GCTGCTGTTG ATAGTGGACG TGGAGGTGTG
 301  GTGAAGACTG CGGCTACTGC GGTGGTGGTG GAGAAGCCGA CGACGGAACG
 351  ATGTCGTTTG AGGTTTTATC AGGGAAAGCC ATTGCCGTTT GGTGCTACTG
 401  CGACAGATGG TGGTGTGAAT TTCGCTGTTT TTCAAGGAAA TGCTACAGCT
 451  GCTACTCTTT GCTTGATCAC TCTTTCCGAT TTACCTGAGA AGAGAGTGAC
 501  CGAGCAAATT TTCCTGGATC CTCTAGCTAA TAAAACTGGA GATGTATGGC
 551  ATGTGTTCCT TAAGGGAGAT TTTGAGAATA TGCTATATGG CTACAAATTT
 601  GATGGGAAAT TCTGTCCTGA AGAAGGACAC TACTTTGACT CTTCGCAGAT
 651  AGTGTTGGAT CCTTATGCCa agGCTATAGT AAGCAGAGGA GAATATGGTG
 701  TATTAGGGCC AGAGGATGAT TGTTGGCCCC CAATGGCTGG CATGGTACCc
 751  TTCTGCTTCT GgATCAGTTT GTATTGGGAA GGAGATCTAC CACTGgAAGT
 801  TTCCcACAgA GAGATCTTGT TnATCnATGA AATGCATGTT CGTGGGTTTA
 851  CTATCCATGA GTCGAGTGAA ACAAAATATC CTGGTACTTA CCTTGGTGTT
 901  GTGGAGAAAC TTGATCACTT GAAGGAACTT GGTGTCAACT GTATAGAGCT
 951  AATGCCCTGT CACGAGTTCA ATGAGCTGGA GTACTATAGT TATAACTCTG
1001  TATTGGGCGA CTACAAGTTT AACTTTTGGG GCTATTCTAC TGTCAATTTC
1051  TTTTCTCCAA TGGGAAGATA CTCATCTGCT GGTCTAAGTA ATTGCGGCCT
1101  CGGTGCAATA AACgAATTTA AGTATCTTGT CAAGGAAGCA CATAAACGTG
1151  GAATCGAgGT TATCATGGAT GTTGTTTTCA ATCACACTGC TGAAGGAAAT
1201  GAAAATGGTC CCATACTATC ATTTAgAgGC ATTGACAACA GTGTGTTTTA
1251  TACGCTAGCT CCTAAGGGTg AATTTTACAA CTACTCaGGA TGTGGAAATA
1301  CCtTCAACTG TAATAATCCC ATTGTACGTC AATtTATAgT GATGCTGAgA
1351  TATtGGGTTA CCGAAATGCA CGTaCATGGC TTCCGCTTTG ATCTTGCTTC
1401  TATCCTTACA AGAAGTAGCA GCTCGTGGAA TGCTGTAAAT GTCTATGGAA
1451  ATTCAATTGA CGGTGACGTG ATCaCCACAG GCACTCCTCT CACAAGCCCA
1501  CCATTGATTG ATATGATTAG CAATGATCCA ATACTTCGTG GAGTAAAGCT
```

1551 TATAGCTGAA GCATGGGATT GTGGAGGCCT TTACCAAGTT GGCATGTTTC

1601 CGCACTGGGG TATCTGGTCG GAGTGGAACG GAAAGTACCG TGACATGGTA

1651 CGGCAGTTCA TCAAAGGCAC TGATGGGTTT TCTGGGGCTT TTGCTGAATG

1701 CCTTTGTGGA AGCCCAAATC TATACCAGAA AGGAGGAAGA AAACCATGGA

1751 ACAGTATAAA TTTCGTGTGT GCCCACGATG GTTTTACTTT GGCTGATTTA

1801 GTGACATACA ACAATAAACA CAATTTGGCA AATGGAGAGG ACAACAAAGA

1851 CGGGGAGAAT CACAATAATA GTTGGAATTG TGGTGAGGAA GGAGAATTTG

1901 CAAGTATCTT TGTGAAGAAA TTGAGGAAAA GACAAATGCG GAACTTCTTC

1951 cTCTGCcTTA tGGTTTCCCA AGGTGTTCCC ATGATATATA TGGGCGATGA

2001 ATATGGTCAC ACTAAGGGAG GAAACaACAA CACGtATTGC CATGATAATT

2051 ATATTAATTA CTTCCGTTGG GATAAGAAGG ATGAATCTTC ATCTGATTTT

2101 TTGAGATTTT GCGGCCTCAT GACCAAATTC CGCCATGAAT GTGAATCACT

2151 GGGATTAGAT GGTTTCCCTA CAGCAGAAAG GCTGCAATGG CATGGTCACA

2201 CTCCTAGAAC TCCAGATTGG TCTGAAACAA GTCGATTCGT TGCATTCACA

2251 CTGGTCGACA AAGTGAAGGG AGAACTATAT ATTGCCTTTA ACGCCAGCCA

2301 TTTGCCTGTA ACGATTACAC TTCCAgATAG GCCTGGTTAT AGATGGCAGC

2351 CGTTTGTGGA CACAGGCAAA CCAGCACCAT TTGACTTCTT GACAGACGAC

2401 GTTCCTGAGA GAGAGACAGC AGCCAAACAA TATTCTCATT TTCTGGACGC

2451 GAACCAGTAT CCGATGCTCA GTTATTCATC CATTATTCTT TTACTATCAT

2501 CTGCTGATGA TGCATAGTTT CATTCACCAA GTTAGGTGGA GGTAAATCAG

2551 CTTCAGATTT TGTTATATGC AGTGAGGTGT TACTTTGTAA ATAAAAGTAA

2601 GAAGCAGGAC AGAACAGAAC TGCAAACGGA TAAAATTTGT GAGGAAGAAG

2651 CTGATGATTT ATAAGAtACA CCTTGTaTTt TAATtGCATT TATATAAAAT

2701 AAAATAnTAG TGAAATTGTc TGTGcGAAAa aaaaaaAAAA AAAAAATAAA

2751 AAAAAAAAAA AAAAAAAAAA AACCATGGTA CCCGGATCCT CGAATTnGAT

2801 ATCAAG

Fig. 11B

Fig. 12


```
   1  CTCATTCCGA GGAATAGACA ACAAGGTTTA TTACATGGTA GATTTGAACA
  51  ACAATGCTCA GCTGCTGAAT TTCGCTGGAT GTGGAAATAC TTTTAACTGC
 101  AATCATCCTA CAGTCATGGA ACTTATACTT GAAAGCTTAA GACACTGGGT
 151  CACCGAGTAT CATGTCGATG GATTTCGCTT TGATCTTGCT AGTGTTCTTT
 201  GCAGAGGGAC AGATGGTACT CCCATTAATG CTCCCCCCCT TGTAAAGGCC
 251  ATTTCCAAAG ATAGTGTATT GTCGAGGTGC AAAATTATTG CTGAGCCATG
 301  GGATTGTGGA GGCCTATATC TTGTTGGAAA GTTTCCGAAC TGGGACCGGT
 351  GGGCTGAGTG GAATGGGAAG TACCGCGATG ACATCAGGAG ATTTATAAAG
 401  GGCGATGCTG GCATGAAAGG AAATTTTGCA ACCCGTATCG CAGGTTCAGC
 451  GGATCTGTAC AGAGTGAACA AGCGAAAGCC GTACCACAGT GTCAACTTCG
 501  TGATTGCCCA TGATGGCTTT ACCTTGTATG ACCTTGTTTC ATACAATAAT
 551  AAGCACAATG ATGCGAACGG TGAAGGTGGC AATGATGGAT GCAATGACAA
 601  CTTCAGTTGG AATTGTGGAA TTGAAGGTGA AACTTCAGAT GCAAATATTA
 651  ACGCACTGCG TTCACGGCAA ATGAAAAATT TTCATTTGGC ACTGATGGTT
 701  TCTCAGGGAA CACCAATGAT GCTTATGGGG GATGAGTATG GGCATACCCG
 751  CTATGGAAAT AATAACAGTT ATGGACATGA TACCGCCATC AACAATTTCC
 801  AGTGGGGACA ATTGGAAGCA AGGAAGAATG ATCACTTCAG GTTCTTTTCC
 851  AAGATGATAA AGTTTCGACT GTCCCACaAt GTTCTTAGAA AGGAaAACTT
 901  CATTGAGAAG AACGACATTA CCTGGCTCGA GGACAACTGG TACAATGAAG
 951  AGAGTAGATT CCTTGCATTT ATGCTCCATG ATGGGAATGG AGGAGATATT
1001  TACTTGGCAT TTAATGCACA CCACTTTTCC ATCAAAACAG CAATACCTTC
1051  ACCACCACGA AATAGAAGTT GGTACCGAGT GGTGGACACT AATCTGAAGT
1101  CACCAGATGA TTTTGTTATT GAGGGAGTGT CTGGTATCAG TGAAACTTAT
1151  GATGTTGCGC CGTACTCTGC TATCCTTCTT GAAGCAAAGC AATAATTACC
1201  GGGACTATGC TGCTTTAGAT GTTGTCCATG TGTTATTACA GTATTACCTC
1251  CTTCTGGATT GGATAGTTCA AATTGGAATT CAGGCTGTTA GCCTATAGAT
1301  GTAGTATGTT GAGCAGAAAT TTTGCAATAA GCAACCAGTT TTGTTCAAAA
1351  AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAA
```

```
  1  LMGLD*EYVI HQVLYLTLL* HCVLLMATSP IQLAVHSRLL SYGSTESTKL
 51  VPSSSGNRGK IVCSLRKLEL EDMNFSGIGR NNDQEAPRRA HRRKALSASR
101  ISLVPSAKRV PTYLFRTDIG GQVKVLVEKT NGKYKVLVEV LPLELSDAHS
151  ELVMVWGLFR SDALCFMPLD LNRRGADGKS STVETPFVQG PSGKVTVELD
201  FEASLAPFYI SFYMKSQLVS DMENSEIRSH RNTNFVVPVG LSSGHPAPLG
251  ISFQPDGSVN FALFSRSARS VVLCLYDDIS VEKPSLEIDL DPYINRSGDI
301  WHAALDCSLP FKTYGYRCKA TTSGKGELVL LDPYAKVIRR VIPRQGGSEI
351  RPKYLGELCL EPGYDWSGDV PPSLPMEKLI IYRLNVTQFT KDKSSKLPDD
401  LAGTFSGISE KWHHFKDLGV NAMLLEPIFP FDEQKGPYFP WHFFSPGNMY
451  GPSGDPLSAI KSMKDMVKKL HANGIEVFLE VVFTHTAEDA PLMNVDNFSY
501  CIKGGQYLNI QNALNCNYPI VQQMILDCLR HWVIEFHIDG FVFVNASSLL
551  RGFNGEILSR PPLVEAIAFD PILSKVKMIA DNWNPLTNDS KENLFPHWRR
601  WAEINMRFCD DIRDFLRGEG LLxNLxTRLC GSGDIFAGGR GPAFSFNYIA
651  RNSGLTLVDL VSFSSNEVAS ELSWNCGQEG ATTNNIVLER RLKQVRNFLF
701  ILFISLGVPV LNMGDECGQS SGGPPAYDAR KSLGWNTLKT GFGTQIAQPI
751  SFLSNLRMRR SDLLQKRTFL KEENIQWHGS DQSPPKWDGP SSKFLAMTLK
801  ADAEVSQTLV SDIVGDLFVA FNGAGDSEIV ILPPPPTDMV WHRLVDTALP
851  FPGFFDEKGT PVEDELVAYE MKSHSCLLFE AQRLAEIDSS KRKKQIRLSS
901  KRQ*PVKPLS IYICLNKRLF FLNK*EDFTE NTCI*TFSFA ASNKKKK
```

Fig. 13

18/19

```
  1  GGRSRTSGSP GLQEFEDPGT MAQSFSISVP HTLDHTLSLP QSSPMELLHC
 51  PSISTYKPKL SFHNHLFSRR SSNGVDFESI WRKSRSSVVN AAVDSGRGGV
101  VKTAATAVVV EKPTTERCRL RFYQGKPLPF GATATDGGVN FAVFQGNATA
151  ATLCLITLSD LPEKRVTEQI FLDPLANKTG DVWHVFLKGD FENMLYGYKF
201  DGKFCPEEGH YFDSSQIVLD PYAKAIVSRG EYGVLGPEDD CWPPMAGMVP
251  FCFWISLYWE GDLPLEVSHR EILxIxEMHV RGFTIHESSE TKYPGTYLGV
301  VEKLDHLKEL GVNCIELMPC HEFNELEYYS YNSVLGDYKF NFWGYSTVNF
351  FSPMGRYSSA GLSNCGLGAI NEFKYLVKEA HKRGIEVIMD VVFNHTAEGN
401  ENGPILSFRG IDNSVFYTLA PKGEFYNYSG CGNTFNCNNP IVRQFIVMLR
451  YWVTEMHVHG FRFDLASILT RSSSSWNAVN VYGNSIDGDV ITTGTPLTSP
501  PLIDMISNDP ILRGVKLIAE AWDCGGLYQV GMFPHWGIWS EWNGKYRDMV
551  RQFIKGTDGF SGAFAECLCG SPNLYQKGGR KPWNSINFVC AHDGFTLADL
601  VTYNNKHNLA NGEDNKDGEN HNNSWNCGEE GEFASIFVKK LRKRQMRNFF
651  LCLMVSQGVP MIYMGDEYGH TKGGNNNTYC HDNYINYFRW DKKDESSSDF
701  LRFCGLMTKF RHECESLGLD GFPTAERLQW HGHTPRTPDW SETSRFVAFT
751  LVDKVKGELY IAFNASHLPV TITLPDRPGY RWQPFVDTGK PAPFDFLTDD
801  VPERETAAKQ YSHFLDANQY PMLSYSSIIL LLSSADDA*F HSPS*VEVNQ
851  LQILLYAVRC YFVNKSKKQD RTELQTDKIC EEEADDL*DT PCILIAFI*N
901  KIxVKLSVRK KKKKKIKKKK KKKKPWYPDP RIxYQ
```

Fig. 14

```
  1  SFRGIDNKVY YMVDLNNNAQ LLNFAGCGNT FNCNHPTVME LILESLRHWV
 51  TEYHVDGFRF DLASVLCRGT DGTPINAPPL VKAISKDSVL SRCKIIAEPW
101  DCGGLYLVGK FPNWDRWAEW NGKYRDDIRR FIKGDAGMKG NFATRIAGSA
151  DLYRVNKRKP YHSVNFVIAH DGFTLYDLVS YNNKHNDANG EGGNDGCNDN
201  FSWNCGIEGE TSDANINALR SRQMKNFHLA LMVSQGTPMM LMGDEYGHTR
251  YGNNNSYGHD TAINNFQWGQ LEARKNDHFR FFSKMIKFRL SHNVLRKENF
301  IEKNDITWLE DNWYNEESRF LAFMLHDGNG GDIYLAFNAH HFSIKTAIPS
351  PPRNRSWYRV VDTNLKSPDD FVIEGVSGIS ETYDVAPYSA ILLEAKQ*LP
401  GLCCFRCCPC VITVLPPSGL DSSNWNSGC* PIDVVC*AEI LQ*ATSFVQK
451  KKKKKKKKKK KK
```

Fig. 15

```
1  GATCATAACT TGAGTTCTAA GCGG
```

Fig. 16A

```
1  CAGGAAACAG CTATGAC
```

Fig. 16B

INTERNATIONAL SEARCH REPORT

Inte Jonal Application No
PCT/GB 98/02280

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/56 C12N15/82 C12N9/44 G01N33/53 C12P21/08
C12Q1/68 C07K16/40 C08B30/00 C08L3/00

According to International Patent Classification(IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C12Q C07K G01N C12P C08B C08L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	WO 96 03513 A (MONSANTO CO) 8 February 1996 see sequence IDs 8 and 9. see page 8, line 18 - line 29 see page 17, line 25 - page 18, line 6	14 1-13, 15-58
Y	WO 95 04826 A (INST GENBIOLOGISCHE FORSCHUNG ;KOSSMANN JENS (DE); EMMERMANN MICHA) 16 February 1995 cited in the application see the whole document	1-13, 15-58
Y	WO 96 19581 A (INST GENBIOLOGISCHE FORSCHUNG ;KOSSMANN JENS (DE); EMMERMANN MICHA) 27 June 1996 see example 6	12

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 November 1998	27/11/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Maddox, A

2

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inte. ional Application No
PCT/GB 98/02280

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No.
X	JAMES M G ET AL: "CHARACTERIZATION OF THE MAIZE GENE SUGARY1, A DETERMINANT OF STARCH COMPOSITION IN KERNELS" PLANT CELL, vol. 7, April 1995, pages 417-429, XP002033602 cited in the application see the whole document -& JAMES, M.G., ET AL.: "Zea mays Sulp (Sugary1) mRNA, partial cds." EMBL ACCESSION NO. U18908,19 April 1995, XP002084161	14
X	US 5 614 619 A (PIEPERSBERG WOLFGANG ET AL) 25 March 1997 see seqeunce ID 2 see table 2	14
X	DATABASE GENESEQ accession no. Q51622, 24 May 1994 "6-o-methyldeoxyguanosine primer A" XP002084445 see abstract & JP 05 273209 A (TAKARA SHUZO) 22 October 1993	14
X	US 5 527 898 A (BAUER HEIDI M ET AL) 18 June 1996 see sequence ID30 see column 73 -& DATABASE GENESEQ Accession no. T44609, 28 January 1997 "Human papillomavirus detection probe MYB187 for HPV type 26" XP002084446 see abstract	14
X	ISHIZAKI Y ET AL: "DEBRANCHING ENZYMES OF POTATO TUBERS (SOLANUM TUBEROSUM L.) I PURIFICATION AND SOME PROPERTIES OF POTATO ISOMYLASE" AGRICULTURAL AND BIOLOGICAL CHEMISTRY, vol. 47, no. 4, 1 January 1983, pages 771-779, XP002001115 see the whole document	36,38
X	WO 92 11382 A (CALGENE INC) 9 July 1992 see the whole document	55

-/--

2

INTERNATIONAL SEARCH REPORT

Inte. Jonal Application No
PCT/GB 98/02280

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No.
X	KORTSTEE, A.J., ET AL.: "Expression of Escherichia coli branching enzyme in tubers of amylose-free transgenic potato leads to an increased branching degree of the amylopectin" THE PLANT JOURNAL, vol. 10, 1996, pages 83-90, XP002084156 see the whole document	55
X	SHEWMAKER C K ET AL: "EXPRESSION OF ESCHERICHIA COLI GLYCOGEN SYNTHASE IN THE TUBERS OF TRANSGENIC POTATOES (SOLANUM TUBEROSUM) RESULTS IN A HIGHLY BRANCHED STARCH" PLANT PHYSIOLOGY, vol. 104, 1994, pages 1159-1166, XP002033871 see the whole document	55
P,X	WO 97 42328 A (PLANTTEC BIOTECHNOLOGIE GMBH ;EMMERMANN MICHAEL (DE); KOSSMANN JEN) 13 November 1997 see sequence ID 1	1-10, 14-16, 22,24, 25,36,38

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Intl. tional Application No
PCT/GB 98/02280

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9603513	A	08-02-1996	AU	691325 B	14-05-1998
			AU	3143795 A	22-02-1996
			CA	2195786 A	08-02-1996
			EP	0772683 A	14-05-1997
			HU	77745 A	28-07-1998
			JP	10504453 T	06-05-1998
			PL	318354 A	09-06-1997
			US	5750876 A	12-05-1998
WO 9504826	A	16-02-1995	DE	4327165 A	16-02-1995
			AU	693469 B	02-07-1998
			AU	7535294 A	28-02-1995
			EP	0713531 A	29-05-1996
			JP	9501052 T	04-02-1997
			CA	2169174 A	16-02-1995
			HU	73740 A	30-09-1996
WO 9619581	A	27-06-1996	DE	4447387 A	27-06-1996
			AU	4433396 A	10-07-1996
			CA	2208410 A	27-06-1996
			EP	0807179 A	19-11-1997
			HU	77470 A	28-05-1998
US 5614619	A	25-03-1997	AU	676461 B	13-03-1997
			AU	2545992 A	27-04-1993
			CA	2119379 A	01-04-1993
			DE	59208758 D	04-09-1997
			DK	604495 T	09-03-1998
			WO	9306219 A	01-04-1993
			EP	0604495 A	06-07-1994
			JP	6510667 T	01-12-1994
			US	5710032 A	20-01-1998
US 5527898	A	18-06-1996	US	5447839 A	05-09-1995
			US	5182377 A	26-01-1993
			US	5639871 A	17-06-1997
			US	5705627 A	06-01-1998
			AT	138108 T	15-06-1996
			AU	645483 B	20-01-1994
			AU	4401189 A	02-04-1990

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter .onal Application No
PCT/GB 98/02280

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5527898 A		CA 1339262 A	12-08-1997
		DE 68926507 D	20-06-1996
		DE 68926507 T	16-01-1997
		EP 0433396 A	26-06-1991
		JP 2651483 B	10-09-1997
		JP 4500910 T	20-02-1992
		WO 9002821 A	22-03-1990
		US 5283171 A	01-02-1994
WO 9211382 A	09-07-1992	US 5349123 A	20-09-1994
		CA 2084079 A	22-06-1992
		EP 0542929 A	26-05-1993
		US 5750875 A	12-05-1998
WO 9742328 A	13-11-1997	DE 19618125 A	13-11-1997
		AU 2891797 A	26-11-1997